FONCTIONS DE TAILLE POUR LA DESCRIPTION ET LA RECONNAISSANCE D'OBJETS $3D$

par

Mohammed Ayoub Alaoui Mhamdi

Thèse présentée au Département d'informatique
en vue de l'obtention du grade de philosophiæ doctor (Ph.D.)

FACULTÉ DES SCIENCES

UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, 18 juillet 2018

Le 13 juillet 2018

Le jury a accepté la thèse de Monsieur Mohammed Ayoub Alaoui Mhamdi dans sa version finale.

Membres du jury

Professeur Djemel Ziou
Directeur de recherche
Département d'informatique

Professeure Marie-Flavie Auclair-Fortier
Membre interne
Département d'informatique

Professeur Hamid Krim
Membre externe
Electrical and Computer Engineering Department
NC State University

Professeur Richard Egli
Président-rapporteur
Département d'informatique

Sommaire

Le but de cette thèse est de concevoir, et de développer des méthodes de reconnaissance d'objets $3D$. Dans ce cadre, nous avons développé deux approches de reconnaissance d'objets $3D$ en nous basant sur leurs points critiques à travers la fonction de taille. La première approche consiste à décrire l'objet $3D$ par un ensemble de fonctions de taille à une seule dimension résultantes de la caractéristique de la concavité et de la convexité pour chaque vertex de l'objet $3D$. La deuxième approche est globale ; dont le principe est la représentation de l'objet $3D$ par une seule fonction de taille. Celle-ci est résultante de la caractéristique du déplacement surfacique et elle est dotée d'invariances aux articulations, aux torsions et aux transformations affines. Afin de pallier le problème de la correspondance partielle d'objets $3D$, une méthode d'apprentissage de la métrique a été utilisée. Nous avons validé les deux approches proposées en utilisant différentes collections d'objets $3D$. Les résultats obtenus se comparent favorablement à ceux proposés dans la littérature.

Mots-clés: Reconnaissance de formes ; classification d'objets $3D$; description d'objets $3D$; correspondance partielle d'objets $3D$.

Remerciements

Je tiens à exprimer ma reconnaissance envers mon directeur de recherche Monsieur Djemel ZIOU Professeur à l'Université de Sherbrooke, qui m'a assuré la direction de ce travail avec beaucoup de compétence et d'efficacité, en faisant toujours preuve à mon égard de compréhension et de gentillesse. Son expérience, son aide, ses encouragements, sa rigueur et ses conseils ont été décisifs dans le développement de ce travail. Je lui exprime toute ma gratitude et ma profonde estime.

J'adresse toute ma gratitude et tous mes remerciements à Madame Marie-Flavie AUCLAIR-FORTIER Professeure à l'Université de Sherbrooke, à Monsieur Monsieur Richard EGLI Professeur à l'Université de Sherbrooke, et à Monsieur Hamid Krim Professeur à l'Université de Caroline du Nord pour avoir accepté d'évaluer ce travail, et aussi d'avoir honoré ce travail au tant que membres de jury.

J'ai le devoir et le plaisir de remercier tous mes collègues au centre de recherche MOIVRE qui m'ont bien accueilli depuis le premier jour de mon arrivée à l'Université de Sherbrooke.

Remerciements

Abréviations

ACP Analyse en composantes principales

CNN Convolutional neural network

CPU Central processing unit

GPU Graphics processing unit

ISOMAP Isometric mapping

PC Personal computer

PCA Principal component analysis

RAM Random-access memory

Table des matières

Liste des figures

Liste des tableaux

Introduction

Durant la dernière décennie, les développements des recherches scientifiques dans les domaines de traitement d'images [1], de vision par ordinateur [2], d'infographie [3] et de synthèse d'images [4] ont permis de consolider le traitement et l'analyse d'objets $3D$ dans différentes applications. En effet, de nombreuses disciplines scientifiques, telles que la géologie [42], l'ingénierie civile [43], l'ingénierie mécanique [44], l'imagerie médicale [45] et l'astronomie [5] reposent sur l'analyse et le traitement de ces données géométriques pour l'avancement de la connaissance. Le traitement de ces données inclut la numérisation [6], la représentation [7], la visualisation [8], la reconnaissance [9], la compression [10], la transmission [11] et le tatouage [12]. L'émergence des images des objets $3D$ est liée directement aux développements des logiciels de la conception assistée par ordinateur [44] et aux avancées dans les technologies d'acquisition telles que les scanners $3D$, les systèmes de stéréovision, et de tomographie. En effet, ils ont favorisé la création de modèles $3D$ dans différents secteurs d'activités humaines, qu'ils soient scientifiques telles que les applications médicales [4], la conception assistée par ordinateur [3], et la gestion du contenu culturel [13], ou ludiques telles que les dessins animés $3D$ [14] et les jeux [15]. Des environnements de programmation pour le traitement des images des objets $3D$ sont disponibles, comme, par exemple OpenGL [16] et Java $3D$ [17]. D'une part, ils ont permis le développement des logiciels de conception assistée par ordinateur afin de faciliter aux professionnels la création, le traitement et l'analyse des images des objets $3D$. D'autre part, ils ont facilité la manipulation des images des objets $3D$ pour les utilisateurs ordinaires à travers la nouvelle génération de téléphones intelligents permettant de visualiser interactivement des modèles $3D$. Dans ce qui suit, nous utilisons les termes objet $3D$ et modèle $3D$ indifféremment pour signifier l'image de l'objet $3D$.

Les procédés d'acquisition d'objets $3D$ diffèrent par les phénomènes physiques mesurables sous-jacents, le contenu informationnel des images acquises et le nombre d'images 2D nécessaires pour capter l'information $3D$. Par ailleurs, selon les applications et les traitements ultérieurs d'objets $3D$, certains procédés permettent l'acquisition des surfaces des objets $3D$ et d'autres les volumes des objets $3D$. Pour cela, nous pouvons distinguer deux approches d'acquisition d'objets $3D$.

Le principe de la première approche d'acquisition d'objets $3D$ est de numériser la forme de l'objet $3D$ sous forme d'un nuage de points $3D$. Les scanners $3D$ [20, 36] et les systèmes de stéréovision [21] sont parmi les procédés d'acquisition d'objets $3D$ dont le résultat de numérisation est un nuage de points $3D$. La représentation de l'objet $3D$ par des nuages de points, fournie d'importantes informations sur la surface de l'objet $3D$. Cependant, pour des besoins de facilité d'estimation des propriétés géométriques, tels que les calculs de la normale et les courbures [22], la représentation par des nuages de points des objets $3D$ est souvent convertie en maillage triangulaire. Avec cette modélisation, l'objet $3D$ est représenté par sa surface frontière composée d'un ensemble de points $3D$ liés par des arêtes afin de former des facettes triangulaires.

Le principe de la deuxième approche d'acquisition d'objets $3D$ est de reconstruire le volume de l'objet $3D$. En tomographie par exemple, l'objet $3D$ est numérisé en un modèle volumétrique sous la forme d'une grille de n^3 voxels, où chaque élément contient une sorte de valeur. Le calcul de cette valeur du voxel est basé sur un ensemble de mesures externes [24]. Elles comprennent souvent l'émission d'un certain signal à travers l'objet vers le capteur et l'analyse de la réponse. La nature des données volumétriques récupérées dépend du capteur. Par exemple, le scanner tomodensitométrique axial consiste à mesurer la capacité du matériel à absorber les rayons X ; ce scanner envoie des faisceaux de rayons X à travers l'objet $3D$, et un capteur mesure leur intensité résiduelle [24]. De même, l'échographie concerne la mesure l'impédance acoustique d'un matériel en mesurant l'atténuation des ultrasons transmis à travers l'objet. Par conséquent, ces deux approches d'acquisition d'objets $3D$ ont proliféré l'archivage des objets $3D$ afin de générer des collections d'objets $3D$. Le but est de les utiliser comme une source d'informations pour l'avancement de la connaissance dans différentes disciplines telles que l'imagerie médicale [25], la sécurité [26] et la reconnaissance faciale [37].

Introduction

La reconnaissance d'objets $3D$ est l'une des capacités inées chez l'être humain [29, 30, 38]. En effet, c'est un processus qui est basé sur l'expérience accumulée [38], puisque l'être humain n'utilise que quelques caractéristiques pour reconnaitre les objets et il les compare avec celles des objets $3D$ qu'il avait déjà rencontrés [38, 41]. Afin d'expliquer cette capacité, des écoles de pensée ont été fondées, dont nous pouvons citer celle de Gestalt [27, 28]. Elle stipule que la reconnaissance d'objets est effectuée par le regroupement des caractéristiques selon plusieurs critères incluant la proximité, la continuité [39] et la symétrie [40]. L'importance de la classification dans le cadre de la reconnaissance d'objets a été étudiée aussi par Mervis et al. [31]. Le principe est lorsque deux ou plusieurs objets sont similaires ou des événements sont traités de manière équivalente, une catégorie existe. Ce traitement équivalent peut s'effectuer selon différents procédés, tels que l'étiquetage des objets similaires ou des événements avec le même nom. Chez l'être humain ce processus ordinaire de reconnaissance d'objets représente une expérience riche. Puisque pour reconnaitre un objet, l'être humain lui associe un nom et une signification selon ses caractéristiques. Pour ce faire, l'être humain fait appel à de nombreux mécanismes et informations dont le contexte permettant de définir l'utilisation ultérieure de l'objet [32, 35] et les relations de l'objet avec les autres objets qui se trouvent dans une scène [33]. En effet, Biederman et al. [34] ont défini cinq relations contextuelles qui lient un objet aux autres objets dans une scène. La première relation contextuelle est l'interposition, qui stipule que tout objet occupe un certain volume. La probabilité est la deuxième relation contextuelle. En effet, elle indique que les objets ont une tendance à se trouver dans certaines scènes mais pas dans d'autres scènes. La position stipule que si l'objet existe dans une scène, il doit se trouver dans une position typique dans la scène. La relation de support exprime que la plupart des objets ne flottent pas. La taille indique que les objets dans une scène ont des tailles relatives entre eux.

Dans cette thèse, nous avons choisi de représenter des caractéristiques des objets $3D$ via des descripteurs qui sont basés sur l'information fournie par leurs points critiques. En effet, les points critiques et les liens entre eux fournissent des informations sur les changements topologiques de la forme [18]. En matière de reconnaissance d'objets $3D$, les points critiques sont associés aux caractéristiques à extraire à travers une fonction définie sur l'objet $3D$. Par conséquent, la comparaison entre les objets $3D$

sera effectuée à travers la mesure de similarité entre les informations fournies par les points critiques associés aux objets $3D$. En effet, le principe de cette description est de représenter une ou plusieurs caractéristiques sous forme d'une fonction mathématique appelée fonction de mesure définie sur l'objet $3D$, et d'encoder quantitativement les propriétés topologiques fournies par les points critiques de cette fonction de mesure à travers la fonction de taille. Le principe de celle-ci est de décrire l'objet $3D$ en calculant le nombre de composantes connexes entre les points critiques de la fonction de mesure définie sur l'objet $3D$. Le résultat est une représentation de l'objet $3D$ sous forme d'un graphe. Le principal avantage de la fonction de taille est le fait qu'elle préserve les invariances dotées par la fonction de mesure. En plus, contrairement aux autres méthodes connexes dans la littérature [46, 47], l'avantage de la description d'objets $3D$ par les fonctions de taille est que la fonction de mesure peut ou ne peut pas être une fonction Morse [18].

Les contributions de cette thèse sont représentées dans deux chapitres. Dans le chapitre 1, une approche locale de la reconnaissance d'objets $3D$ à travers un ensemble de fonctions de taille est décrite. Le principe est de découper l'objet $3D$ en un ensemble de 18 portions ; afin de représenter chacune par une fonction de taille résultante des caractéristiques de concavité et de convexité sur chaque point de la portion $3D$. Un processus de normalisation est effectué afin de doter la méthode l'invariance aux articulations, torsions et transformations affines. Nous avons validé notre approche en utilisant différentes collections d'objets $3D$.

Afin de surpasser les limitations de la contribution décrite dans le chapitre 1, dans le chapitre 2, nous proposons une méthode globale de la reconnaissance d'objets $3D$ par une seule fonction de taille. En effet, cette dernière est résultante de la caractéristique du déplacement surfacique et elle est dotée d'invariance aux articulations, aux torsions et aux transformations affines. En plus, afin de pallier au problème de la correspondance partielle d'objets $3D$, un simple réseau de neurones est utilisé pour l'apprentissage de la métrique. Nous avons validé notre méthode en utilisant différentes collections d'objets $3D$ complets et une collection d'objets $3D$ avec des parties manquantes.

La dernière partie, présente les conclusions et les perspectives relatives aux contributions réalisées.

Chapitre 1

Une approche locale de la reconnaissance d'objets $3D$ à travers un ensemble de fonctions de taille

Dans ce chapitre, nous présentons une approche locale de la reconnaissance d'objets $3D$ à travers un ensemble de fonctions de taille. Le principe est de morceler l'objet $3D$ en un ensemble de portions et décrire chacune par une fonction de taille résultante d'une fonction de mesure représentant les caractéristiques de convexité et de concavité. Afin de doter la méthode proposée de l'invariance aux articulations, aux torsions et aux transformations affines, un processus de normalisation est effectué en utilisant la projection ISOMAP pour que les objets $3D$ de la même classe aient les mêmes articulations et torsions ; et en utilisant l'ACP pour que les objets $3D$ de la même classe aient la même position, orientation et échelle. Nous avons validé la méthode proposée en utilisant la collection d'objets $3D$ de McGill composée de 457 objets $3D$ et une collection étendue d'objets $3D$ composée de 2902 modèles $3D$. Les scores obtenus sont favorablement comparables avec ceux obtenus dans l'état de l'art.

J'ai réalisé la totalité de ce travail sous la supervision du Professeur Djemel Ziou. Nous présentons notre approche locale de la reconnaissance d'objets $3D$ dans les pages

qui suivent, dans un articule intitulé **A local approach for** 3D **object recognition through a set of size functions**. Cet article a été publié dans la revue **Image and Vision Computing, Elsevier** en 2014. La version manuscrite de cet article a été modifiée suite aux recommandations des membres de jury et cela n'engage que ma responsabilité.

A local approach for $3D$ object recognition through a set of size functions

Mohammed Ayoub Alaoui Mhamdi
Département d'informatique, Université de Sherbrooke,
Sherbrooke, Québec, Canada J1K 2R1
ayoub.alaoui@usherbrooke.ca

Djemel Ziou
Département d'informatique, Université de Sherbrooke,
Sherbrooke, Québec, Canada J1K 2R1
djemel.ziou@usherbrooke.ca

Keywords: $3D$ model description, Object recognition, Object categorization, Shape classification.

Abstract

In this paper, a local approach for $3D$ object recognition is presented. It is based on the topological invariants provided by the critical points of the $3D$ object. The critical points and the links between them are represented by a set of size functions obtained after splitting the $3D$ object into portions. A suitable similarity measure is used to compare the sets of size functions associated with the $3D$ objects. In order to validate our approach's recognition performance, we used different collections of $3D$ objects. The obtained scores are favourably comparable to the related work.

1.1 Introduction

The recognition of 3D objects has been widely studied during the last decade [29, 30]. Its applications are various, including medical imaging [31, 32], virtual reality [33], face detection [34, 35], cultural heritage [36], and quality control [37]. The principle is to recognize 3D objects based on their shape properties extracted by the mean of a shape descriptor. The recognition process has three major steps. The first step is the computation of descriptors, which can be global [1], local [2], structural [11, 12, 53], transform-based [6, 59], or view-based [7, 52]. The second step is the similarity measure which can be distance-based [17], probability-based [18], or graph-based [48]. A decision-making process is performed in the third step in order to classify the 3D objects according to the similarity measure used, and it can be hard [29, 30], or soft [30, 50].

In our work, we chose to take advantage of the topological invariants provided by the critical points of the 3D object and the link between them. In fact, the critical points and the links between them provide the topological changes of the 3D object through a function defined on it. The topological changes have been previously used for 2D and 3D object recognition [8–12,16,54]. The main idea behind the definition of the size function consists in encoding quantitatively these topological changes. In fact, it relies on a real continuous function defined on the shape called measure function [11], which is chosen to capture some specific features of the 3D object [10].

This paper is organized as follows : in Section 2, the related work is presented and discussed. Our proposed method is detailed in Section 3. The pre-processing stage of the 3D objects is detailed in Section 4. The interpretation of the obtained results using the proposed method and some well-known methods is presented in Section 5. Finally, conclusions and future work are presented in Section 6.

1.2 Related work

The 3D object recognition can be implemented by the comparison of a 3D object to several objects, several objects to a model, or several objects to several models [50]. The decision can be hard [29, 30], in this case the 3D object can be described

or not by one of the $3D$ models. It can be soft [30] ; in this case, the $3D$ models are sorted according to their resemblance to the $3D$ object to be recognized. Indeed, there are several strategies to identify similar objects such as the nearest neighbours [50], the ranking [50], the Bayesian decision rule [60], the risk functions [67], and the discrimination [68].

The criterion of similarity and decision are strongly related to the feature space and to the $3D$ object collection's size [50, 51]. The $3D$ shape descriptors can be regrouped into five categories [20, 21] which are global, local, transform-based, views-based, and structural. In what follows, we present the different groups of approaches, and we indicate for each one of them the description principle, the similarity measure, and the used decision.

In the global approaches, the description of the $3D$ object relies on the measure of its geometric properties. It can be performed by calculating, for example, distances between vertices, or by a histogram associated with a specific feature of the $3D$ object. Ion et al. [1] described the $3D$ object by an eccentricity histogram of its vertices. The similarity measure between the $3D$ objects is performed by using the Euclidean distance between their associated histograms, and the decision is soft. Rabin et al. [3] characterized the $3D$ object by a global descriptor that contains the geodesic information of its mesh. At first, the $3D$ object is sampled to 500 vertices. Afterwards, for each mesh vertex representing the $3D$ object, the geodesic distance, the median geodesic distance, and the eccentricity transforms are computed. The global descriptor associated with the $3D$ object is composed of the geodesic information of the 500 vertices. Finally, the similarity measure between the $3D$ objects is performed by using the distance of Wasserstein [61] between their associated global descriptors, and the decision is soft.

In the local approaches, the description of the $3D$ object relies on the neighborhood of each vertex on it. Assflag et al. [2] developed the idea that the shape of a $3D$ object can be described by its surface curvatures map. The description of the $3D$ object is performed in four steps. The first step consists in smoothing and decimating the vertices of the $3D$ object. During the second step, the principal curvatures are estimated on each vertex. The third step consists in deforming the surface of the $3D$ object to adapt it to a surface of a sphere and keep up the curvatures of the

initial surface. The mesh's surface obtained after the deformation is projected to an image which encodes the vertex's position and curvatures of the original surface in the fourth stage. The obtained image represents the curvatures map of the $3D$ object. The similarity measure between the $3D$ objects is performed by using the distance of Kolmogorov-Smirnov between their associated curvatures map, and the decision is soft. Xiaofeng et al. [71] described the $3D$ object by the projected areas of its mesh's vertices. At first, for each vertex of the mesh that represents the $3D$ object, the projected area on the vertical plane of the normal vector is computed. Afterwards, the result is transferred to the $1D$ Fourier transform to obtain the integral invariants from which the feature vector is formed. Finally, the similarity measure between the $3D$ objects is performed by using the Euclidean distance between their associated feature vectors, and the decision is soft.

The idea of the transform-based approaches is that in the transform domain, some coefficients of high frequency contain most of the energy, especially when the transform is orthogonal. For this purpose, the description of the $3D$ object relies upon the use of these coefficients as a feature vector. Vranic et al. [6] used the coefficients of the Fourier transform like a feature vector which represents the $3D$ object. At first, the $3D$ object is centered, scaled, and aligned by the principal components analysis [15] in order to ensure invariance to translation, scale, and orientation. Afterwards, the $3D$ object is discretized in a grid of voxels. Finally, the first 172 coefficients of Fourier are computed, which form the feature vector of the $3D$ object. The similarity measure between the $3D$ objects is performed by using the Euclidean distance between their associated feature vectors, and the decision is soft. Laga et al. [59] developed three $3D$ object descriptors extracted from the spherical wavelet. In the two other descriptors, the feature vectors of the $3D$ object are formed respectively from the L_1, and L_2 energies of the spherical wavelet sub-bands. At first, the $3D$ object is uniformly sampled using a geodesic dome. Afterwards, the three descriptors associated with the $3D$ object are computed. Finally, the similarity measure between the $3D$ objects is performed by using the $L_2 - distance$ between their associated feature vectors, and the decision is soft.

In the view-based approaches, the $3D$ object is represented by a set of $2D$ views, where each one is described by a $2D$ descriptor. Vranic [7] described a $3D$ object by six

depth images. In order to ensure invariance to translation, to scale, and to orientation, the $3D$ object is normalized by the principal components analysis [15]. Subsequently, the $3D$ object is set within its bounding box from which the six gray level depth images are computed and transformed in the Fourier domain. The feature vector is composed of the absolute value of the Fourier coefficients of the six depth images. The similarity measure between the $3D$ objects is performed by using the Euclidean distance of their associated feature vectors, and the decision is soft. Mahmoudi et al. [56] described the $3D$ object by seven characteristics $2D$ views. Each $2D$ view is represented by the curvature scale space descriptor of its $2D$ shape [57]. For the similarity measure between the $3D$ objects, the authors used the geodesic distance defined in [66] in order to compare the curvature scale space descriptors of their associated $2D$ views, and the decision is soft.

In the structural approaches, the description of the $3D$ object relies on the analysis of its topological changes based on a function defined on it. The size function has been used for $3D$ objects matching for the first time in the work of Biasotti et al. [11]. At first, they described the $3D$ object by a suitable skeletal graph based on the construction of a centreline skeleton. Afterwards, they defined the skeletal graph by a set of measure functions, which captures quantitative attributes of the $3D$ object, and they computed the associated size functions with the help of D'Amico's algorithm [23]. A suitable distance between size functions is used to perform the similarity measure between the $3D$ objects, and the decision is soft. In another work, multidimensional size functions are introduced [12] to describe the $3D$ object defined by either triangular meshes or by voxels volume. The idea is that the shape of a $3D$ object can be described by a set of measuring functions, each one representing some specific features of the $3D$ object. In order to reuse the $1D$ size functions' principles, they conceived a theorem which generalizes the $1D$ size functions' concepts for the multidimensional case. For the similarity measure, they developed a lower bound matching distance to compare the multidimensional size functions associated with the $3D$ objects, and the decision is soft.

Table 1.1 represents a summary of the mentioned methods for describing $3D$ objects, which belong to the five approaches of $3D$ object recognition. From Table 1.1, we noticed that most of the description methods rely on the triangular mesh that

tableau 1.1 – Summary of methods for recognizing 3D objects.

Approaches	References	Descriptors	Types of representation	Invariances	Similarity measures	Decision	Collection of 3D objects	Computer configuration
Global	[1]	Eccentricity histogram	Mesh and volume	Affine transforms, articulation, and torsion	Euclidean distance	Soft	255 3D objects of the McGill Shape Benchmark	-
	[3]	Global descriptor of geodesic information	Mesh	Affine transforms, articulation, and torsion	Distance of Wasserstein	Soft	457 3D objects of the McGill Shape Benchmark	-
Local	[2]	Curvatures map	Mesh	Affine transforms	distance of Kolmogorov-Smirnov	Soft	250 3D models from the web and CAD	-
	[71]	Projected areas of vertices	Mesh	Affine transforms	Euclidean distance	Soft	867 engineering models of ESB database	-
Transform-based	[6]	3D Fourier descriptor	Mesh	No	Euclidean distance	Soft	1830 3D objects from 3Dcafe and viewpoint sites	850 MHz Processor Pentium III, windows 2000
	[59]	Spherical wavelet coefficients, L_1, and L_2 energies of the spherical wavelet sub-bands	Mesh	No	Euclidean distance	Soft	1814 3D objects of the Princeton Shape Benchmark	-
View-based	[7]	Fourier descriptor for the six depth images of the 3D object	Mesh	No	Euclidean distance	Soft	1841 3D objects from 3Dcafe and viewpoint sites	-
	[57]	Curvature scale space descriptors associated to the seven 2D views	Mesh	No	Geodesic distance	Soft	100 models of the Mpeg 7 database	-
Structural	[11]	Size functions	Mesh and volume	Affine transforms, articulation, and torsion	Matching distance	Soft	CAD models of Drexel University, Princeton Shape Benchmark, McGill Shape Benchmark, and humanoid of Cesar	1.73 GHz laptop PC-M
	[12]	Multi-dimensional size functions	Mesh, and volume	Affine transforms, articulation, and torsion	Lower bound matching distance	Soft	255 3D objects of the McGill Shape Benchmark	AMD Athlon 3500, 2 GB RAM

represents the 3D object. The usage choice of descriptors presented in Table 1.1 is related to the collection of the 3D objects, which categories may contain some specific problems. For example, structural descriptors [11, 12] and global descriptors [1, 3] showed invariance to affine transforms, to articulations, and to torsions when they

were validated using the articulated $3D$ models of the $3D$ object collection of McGill [19]. We noticed that some of these descriptors [1, 11, 12] were validated by using only 255 $3D$ models of the $3D$ object collection of McGill. Meanwhile, the description process for all the mentioned methods did not rely on the contextual information of the $3D$ objects, because the used collections of $3D$ objects contain isolated $3D$ models. For the similarity measure, most of the description methods used the Euclidean distance in order to compare between the $3D$ objects through their associated feature vectors. However, the structural descriptors rely on the correspondence between the graph-based features representing the $3D$ objects as a similarity measure. Regarding the decision, the retrieval process is performed by using a soft decision for all the mentioned descriptors.

Unlike the first work of Biasotti et al. [11], which requires the definition of the triangular mesh by skeletal graph before the descriptions by the set of size functions, and the second work of Biasotti et al. [12], where the description of the $3D$ object relies on a multidimensional size function resulting from a multidimensional measure function defined on its whole, our method takes advantage of the different spatial information provided by each benchmark axis. For that purpose, the $3D$ object is split into a set of portions. Each portion is described by a size function resulting from a measure function, which is chosen to capture the concavities and the convexities features of the $3D$ object. In the next section, we will go further into the principles of size functions in the context of $3D$ object recognition.

1.3 Proposed method

The size function has been developed since the beginning of the 1990s by Frosini et al. [69] and Verri et al. [70]. It was used by Handouyahia et al. [9] for the recognition of sign language, by Ferri et al. [4] for the automatic classification of white blood cells, by D'Amico et al. [45] for the automatic classification of melanocytic lesions, and by Biasotti et al. [11, 12] for the $3D$ object recognition. Its principle is to describe the $3D$ object by encoding the topological changes provided by its critical points and the links between them. In fact, the critical points are those of a real function defined on the $3D$ object, called measure function. The values taken by the measure function on

its critical points are called critical values. A critical point can be a local or a global maximum or minimum or saddle point of the measure function defined on the $3D$ object. The idea is to describe the $3D$ object by a feature as a function. For example, the measure function can be the distance between the center of the mass of the $3D$ object and another of its points. The critical points of this function may have physical significations, such as the appearance or disappearance of one or several connected components. The computation of critical points is performed by a topological analysis of the measure function as it will be explained in Section 3.1. To illustrate, we consider the cyclide in Figure 1.1-(a) defined by a function φ which takes values in the R axis. The critical points P_1, P_2, P_3, and P_4 associated to the critical values a, b, c, and d, indicate the appearance or the disappearance of one or several connected components. For example, between the critical values a and b associated to the critical points P_1 and P_2, there is an only one connected component, between the critical values b and c associated to the critical points P_2 and P_3, there are two connected components, and so on.

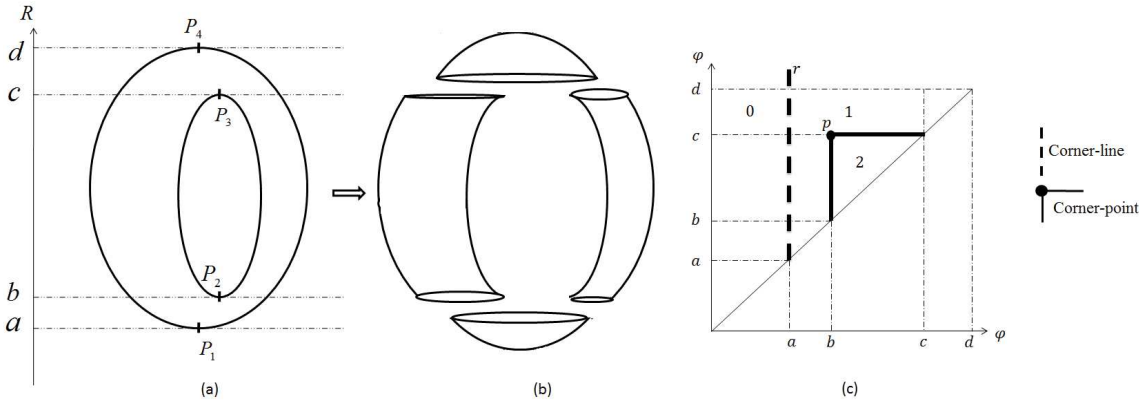

figure 1.1 – (a) The critical points, (b) the connected components, (c) and the resulting size function.

However, if we associate with the same cyclide a measure function ψ that takes values in the R' axis, it will be described by other critical points, and consequently by other connected components as illustrated in Figure 1.2. Therefore, the definition of different measure functions in the same $3D$ object leads to different descriptions.

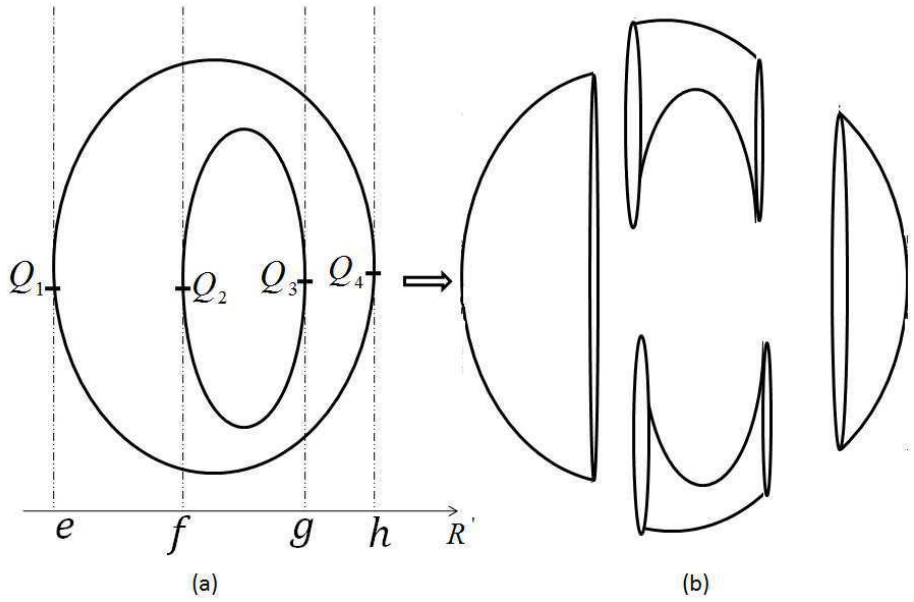

figure 1.2 – (a) The critical points, (b) and the connected components according to the R' axis.

1.3.1 $3D$ object description

Let M be a connected topological space endowed by a continuous function φ defined as follows : $\varphi : M \to \mathbb{R}$. The pair (φ, M) is called a size pair, and φ is the measure function. Let $l_{(M,\varphi)} : \mathbb{R} \times \mathbb{R} \to \mathbb{N} \cup \{\infty\}$ defined by $l_{(M,\varphi)}(x, y)$ the size function corresponding to the number of connected components of the subset $\{P \in M, x < \varphi(P) < y\}$. Consider the example of the Figure 1.1-(a). When $a < \varphi(P) < b$, there is only one connected component, the size function takes the value one. When $b < \varphi(P) < c$, there are two connected components, the size function takes the value two, and so on. The result is the size function graph represented in the Figure 1.1-(c), in which the regions where the size function takes different constants are delimited by corner-points. In the Figure 1.1-(c), the bold dashed line r is a corner-line, and it indicates the occurrence of the first critical point associated to the minimal critical value taken by the measure function φ. The point p is a corner-point that represents the intersection point between the two bold solid lines, which delimits the region where the size function takes the value one, and the region where the size function

φ takes the value two. In fact, the corner-point $p = (b, c)$ represents a detail of the $3D$ object, and it is associated with the local minimum critical point P_2 where the measure function takes the critical value b, where two new connected components appear, and the saddle critical point P_3 where the measure function takes the critical value c, where the two connected components disappear.

The aim is to represent the $3D$ object by a measure function on its mesh, which describes the concavities and convexities features. However, one measure function may not be sufficient to extract relevant spatial information of the $3D$ object. A feasible scheme is to segment the $3D$ object into a set of portions, and to represent each one of them by a size function. However, the existing $3D$ object segmentation algorithms [72, 73, 74] have some limitations. In fact, if the $3D$ object underwent a high smoothing (resp. noise) rate, it would be under-segmented (resp. over-segmented). Henceforward, in our method we chose to take advantage of the different spatial information on each benchmark axis, by splitting the $3D$ object into 18 portions ; as illustrated in the example of Figure 1.3. For example, for the case of $x \geq 0$, only the vertices that have positive abscissas are selected ; the others are omitted from the mesh data structure, and consequently ; the referenced faces are updated, and so on for the other cases. Afterwards, each portion is represented by a measure function that describes the concavities and convexities features. It can be expressed by the function $\varphi(P_i) = e^{-d_{Euclidean}(O,P_i)}$, where $d_{Euclidean}(O, P_i)$ is the Euclidean distance of each vertex P_i in the portion from the benchmark center O. In other words, the use of an exponent in the measure function, defines how each portion's vertex is viewed from the benchmark center O. According to the portion's topology, a vertex can belong to a convex or to a concave region, which expresses a topological change that will be encoded by the resulting size function. For that purpose, for each portion, the measure function $\varphi(P_i) = e^{-d_{Euclidean}(O,P_i)}$ is computed on each vertex P_i, and its associated graph is constructed by finding the adjacent vertices of P_i. However, the measure function's graph has often a large number of vertices, which requires a large storage capacity in memory, and makes the computation of size function more time-consuming. To overcome these limitations, the portion of the $3D$ object can be sampled in a small set of vertices, before computing the measure function. Meanwhile, the sampling process of the portion may cause the loss of some details of

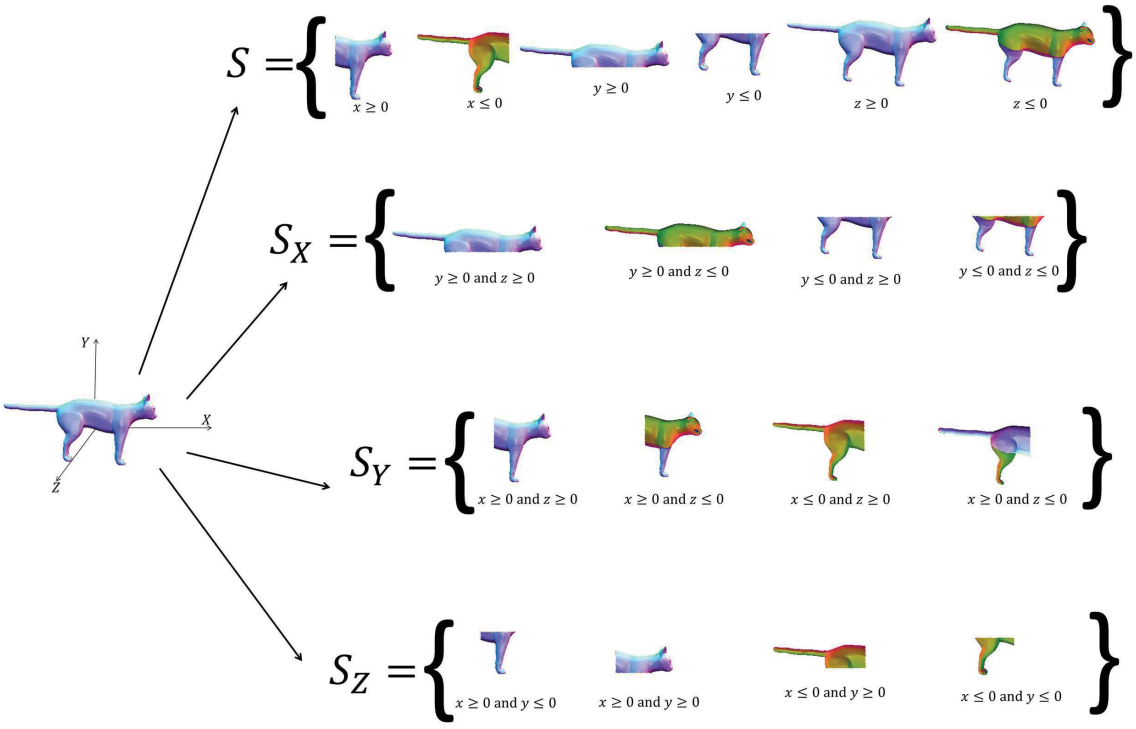

figure 1.3 – Splitting a $3D$ object into 18 portions.

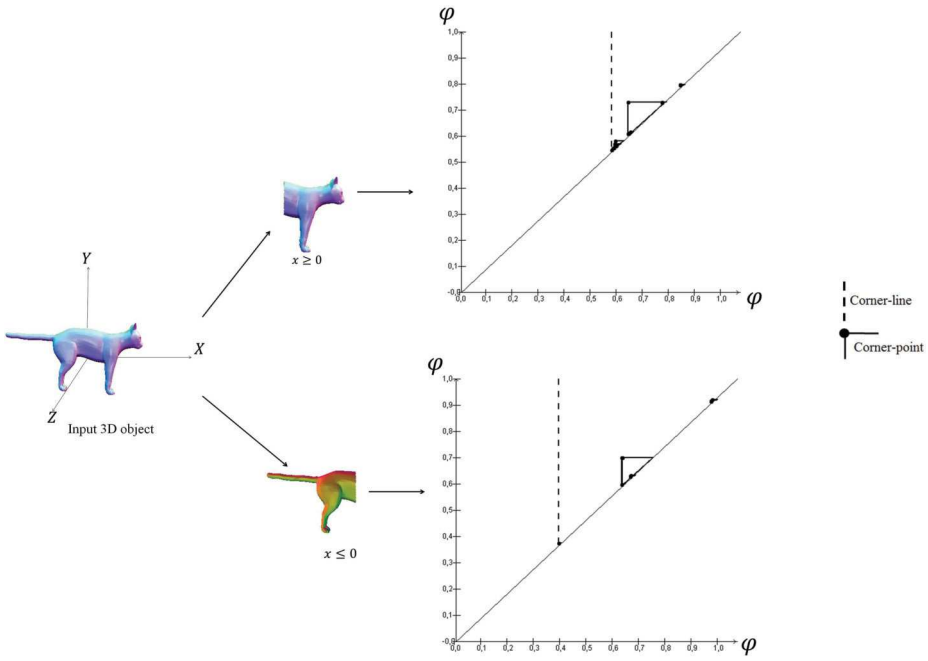

figure 1.4 – Example of size functions associated to two portions.

it that can be relevant for its description. In order to avoid this problem, the measure function's graph is simplified through a $\Delta^* - reduction$ process [23, 78]. In fact, it was demonstrated by Frosini et al. [78] and D'Amico [28], that this reduction process of the measure function's graph does not influence the resulting size function. The aim of simplifying the measure function's graph is to obtain an oriented arborescence that contains only the critical points and the links between them. This arborescence is used for the direct computation of corner-lines and corner-points of the resulting size function with the help of D'Amico's algorithm [23]. Finally, each $3D$ object will be described by a set of 18 size functions. Figure 1.4 illustrates the description of two portions of a $3D$ object by size functions.

However, the splitting of the same $3D$ object, but with a different position, scale, orientation, articulation, or torsion can generate different portions, and consequently, different descriptions by the set of 18 size functions. In order to overcome this invariance problem, the $3D$ object must undergo a preprocessing stage, which will be detailed in section 4.

1.3.2 Similarity measure

As mentioned previously, each $3D$ object is described by a set of 18 size functions. Thus, the comparison between two $3D$ objects is performed through the similarity measure between their associated set of 18 size functions. However, this similarity measure can be expressed in several manners. It can be expressed by finding the minimum distance among the matching distances between the 18 size functions associated with the $3D$ objects, but it can create confusion between $3D$ objects of different categories. Since we can find $3D$ objects belonging to a category that have one or several similar portions to those of $3D$ objects that belong to another category. For example, the ants and the spiders have almost the same abdomen, and consequently, the description by the size function of the ant's abdomen and description by the size function of the spider's abdomen will be almost the same. In order to avoid this drawback, one of the possible solution is to express the similarity measure as an average of minimum distances between their associated 18 size functions. In fact, it was demonstrated by Terrades et al. [76] that the average of minimum distances got better results than the minimum of minimum distances. The principle is that if a $3D$ object is very similar to the $3D$ query object, the average of minimum distances between its associated set of 18 size functions, and those of the query object is the most minimal, comparing to those of the other $3D$ objects. Meanwhile, according to Figure 1.4, six portions of the $3D$ object are extracted by relying on one constraint, and their associated size functions are stored in the set S, and there are twelve portions that are extracted from the $3D$ object by relying on two constraints but according to the different benchmark axis. The subsets S_X, S_Y, and S_Z represent, respectively the portions according to the X, Y, and Z axis that were extracted by relying on two constraints, and each one of them contains four size functions. For that purpose, we chose to ramify the average of the minimum distances between 18 size functions associated to the $3D$ objects into two parts. The first part, expresses the average of the minimum distances between the size functions associated with the portions that were extracted by relying on one constraint. The second part, expresses the average of the average minimum between the twelve remaining size functions, and they are associated with the portions that were extracted by relying on two constraints. Each 3D object Γ is represented by the set of size functions $F_s^\Gamma = \{S^\Gamma, S_X^\Gamma, S_Y^\Gamma, S_Z^\Gamma\}$, with

$S^\Gamma = \{S_i^\Gamma\}_{(i=1,\ldots,6)}$, $S_X^\Gamma = \{S_{X_i}^\Gamma\}_{(i=1,\ldots,4)}$, $S_Y^\Gamma = \{S_{Y_i}^\Gamma\}_{(i=1,\ldots,4)}$. The first part of the similarity measure between two $3D$ objects A and B is formulated as follows :

$$d_1(A,B) = \frac{1}{6}\sum_{i=1}^{6} d_{min_1}(S_i^A, S^B) \tag{1.1}$$

Where $d_{min_1}(S_i^A, S^B)$ is the minimum distance between the i^{th} size function S_i^A of the subset S^A associated with the $3D$ object A and all the size functions of the subset S^B associated to the $3D$ object B, and it can be defined as follows :

$$d_{min_1}(S_i^A, S^B) = \min_{1\leq j\leq 6} d_{match}\left(S_i^A, S_j^B\right) \tag{1.2}$$

Where $d_{match}\left(S_i^A, S_j^B\right)$ is the matching distance between the i^{th} size function S_i^A of the subset S^A associated to the $3D$ object A and the j^{th} size function S_j^B of the subset S^B associated to the $3D$ object B. The size function is represented by a graph composed of a finite number of corner-lines and corner-points in the upper diagonal. For that purpose, the similarity measure between two size functions is graph-based, and it can be performed by computing the correspondence between their associated graphs. Let two size functions l_1 and l_2, with their associated sets of corner-lines and corner-points C_1 and C_2 respectively. The principle is to measure the cost of displacing the corner-lines and corner-points of l_1 to those of l_2 as shown in Figure 1.5. The matching distance between l_1 and l_2 is defined in [13] as follows :

$$d_{match}\left(l_1, l_2\right) = \min_{\sigma}\max_{p\in C_1}\delta\left(p, \sigma(p)\right) \tag{1.3}$$

Where σ varies among all the bijections between C_1 and C_2. In fact, the corner-lines and the corner-points of l_1 and l_2 are stored respectively in the sets C_1 and C_2 in the ascending order of their abscissas. In the examples of Figure 1.5, the minimum constraint mentioned in the equation (3) for the bijection σ between the sets C_1 and C_2, implies that the matching between l_1 and l_2 is given by the displacement of r to r', a to a', b to b', and c to Δ_c. In general, if the size function l_2 has an inferior number of corner-lines and corner-points than those of the size function l_1, its set of corner-points is completed by the orthogonal projection of the remaining corner-

points of the size function l_1 at the diagonal, as illustrated in the example of Figure 1.5-(c). δ is the displacement cost of a corner-point $p = (x, y)$ that belongs to the set of corner-points C_1 associated with the size function l_1 ; to a corner-point $p' = (x', y')$ that belongs to the set of corner-points C_2 associated with the size function l_2, and it is defined [13] as follows :

$$\delta\left(p, p'\right) = \min\left\{\max\{|x - x'|, |y - y'|\}, \max\left\{\frac{|y - x|}{2}, \frac{|y' - x'|}{2}\right\}\right\} \tag{1.4}$$

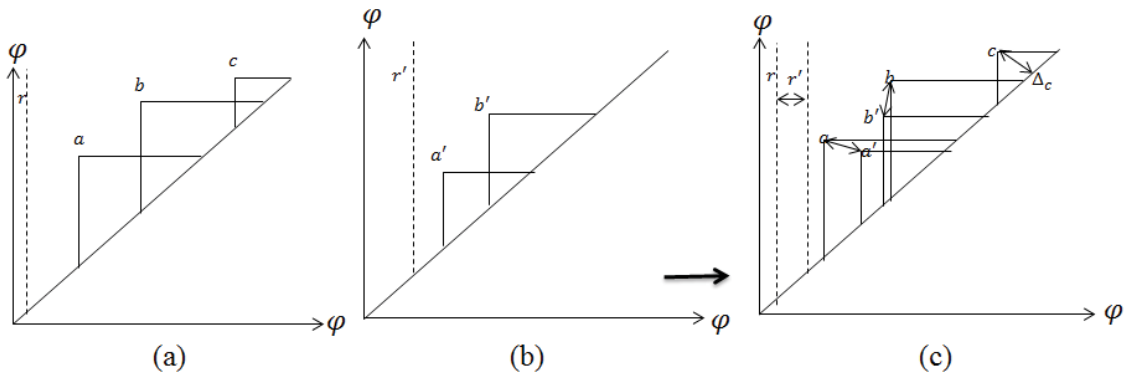

figure 1.5 – Example of matching between two size functions (a) l_1 with $C_1 = \{r, a, b, c\}$ and the size function (b) l_2 with $C_2 = \{r', a', b'\}$, given by $d_{match}(l_1, l_2) = \max\{\delta(r, r'), \delta(a, a'), \delta(b, b'), \delta(c, \Delta_c)\}$.

The second part of the similarity measure expresses the average of the average minimum between the twelve remaining size functions. For that purpose, it is ramified into three subparts expressed as follows :

$$d_X(S_X^A, S_X^B) = \frac{1}{4}\sum_{i=1}^{4} d_{min_2}(S_{Xi}^A, S_X^B), \tag{1.5}$$

Where $d_{min_2}(S_{X_i}^A, S_X^B) = \min_{1\leq j\leq 4} d_{match}\left(S_{X_i}^A, S_{X_j}^B\right)$.

$$d_Y(S_Y^A, S_Y^B) = \frac{1}{4}\sum_{i=1}^{4} d_{min_2}(S_{Yi}^A, S_Y^B), \tag{1.6}$$

$$d_Z(S_Z^A, S_Z^B) = \frac{1}{4}\sum_{i=1}^{4} d_{min_2}(S_{Zi}^A, S_Z^B). \tag{1.7}$$

The second part of the similarity measure can be formulated as follows :

$$d_2(A, B) = \frac{d_X(S_X^A, S_X^B) + d_Y(S_Y^A, S_Y^B) + d_Z(S_Z^A, S_Z^B)}{3} \tag{1.8}$$

Finally, the similarity measure between the two $3D$ objects A and B is formulated as follows :

$$d(A, B) = \frac{d_1(A, B) + d_2(A, B)}{2} \tag{1.9}$$

1.4 $3D$ object preprocessing

The available digital $3D$ objects in the various $3D$ objects collections are generated from $3D$ acquisition devices such as $3D$ scanners [38], and stereovision systems [39, 75, 77] or by the computer aided design software [46, 47]. However, these processes have some limitations which can be mechanical, optical, and in accuracy. And consequently, they often generate meshes with artefacts like representations with gaps and overlapping regions as illustrated in the examples in Figure 1.6. Thus, the presence of artefacts in the mesh influences the features to be extracted from the $3D$ objects. Therefore, the $3D$ objects must undergo a preprocessing stage, which can be performed in two steps. The first step is a low level preprocessing, which consists in repairing the mesh of the $3D$ object by removing artefacts to ensure its connectivity, and to enhance its quality for further usage. The second step consists in normalizing the $3D$ object in order to endow the features to be extracted invariance to affine transforms, articulations, and torsions. In what follows, we will describe the principle of each step.

1.4.1 Mesh repairing

This step consists in repairing the mesh of the $3D$ object by removing its artefacts like holes and separating or overlapping regions. The main idea consists in the approximation of the mesh in a way that the resulting one does not contain artefacts. For

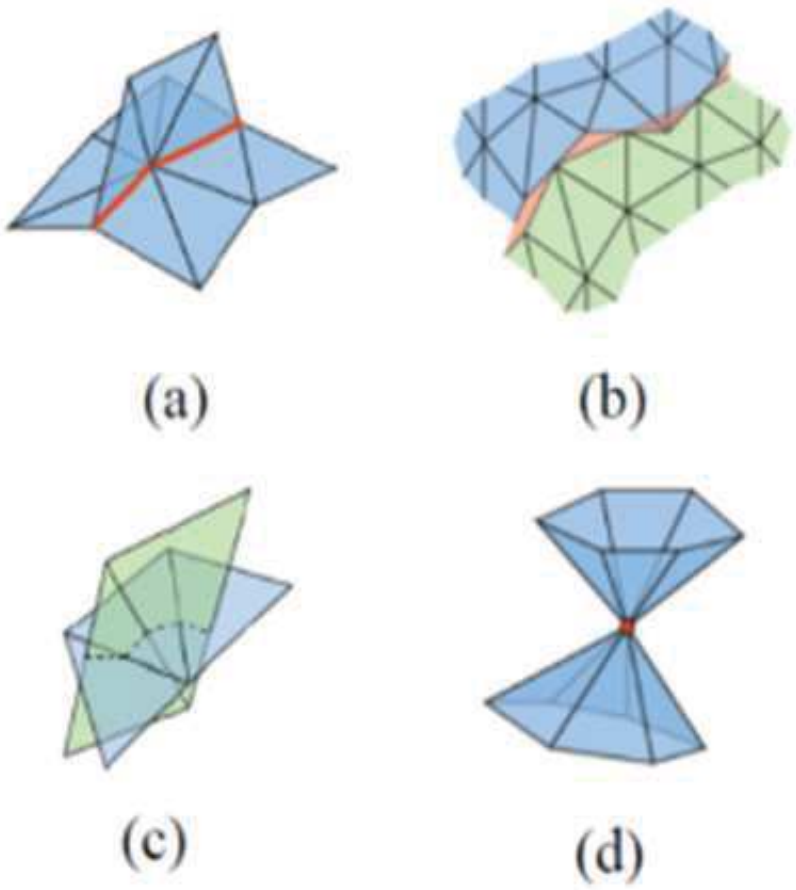

figure 1.6 – Example of artefacts : (a) complex edges, (b) separated regions, (c) overlapping regions, (d) and a singleton vertex (Examples taken from [28]).

this purpose, there are two approaches of mesh repairing. The first approach is surface oriented, where the repairing process relies on filling holes, and stitching separated or overlapping regions. However, the available algorithms [40, 41, 42, 63] belonging to this approach cannot resolve all the artefacts, especially when all the mesh triangles are isolated. Henceforward, we chose to proceed with the second approach which is volume oriented. The principle consists in converting the input defective mesh into an intermediate volumetric representation from which the corrected mesh is extracted. For that purpose, the $3D$ object is discretized under GPU in a grid of 256^3 voxels. At first, the $3D$ object is placed inside its bounding cube. This latter is subdivided regularly into of 256^3 voxels. Afterwards, each voxel that overlaps with one of the mesh triangles is marked to determine the boundary of the $3D$ model in the grid of voxels. The overlapping test is performed with the help of Akenine-Möller's algorithm [27]. Finally, the obtained grid has 256^3 voxels, from which the corrected mesh is extracted by using the marching cubes algorithm [5]. Figure 1.7 illustrates the mesh repairing process.

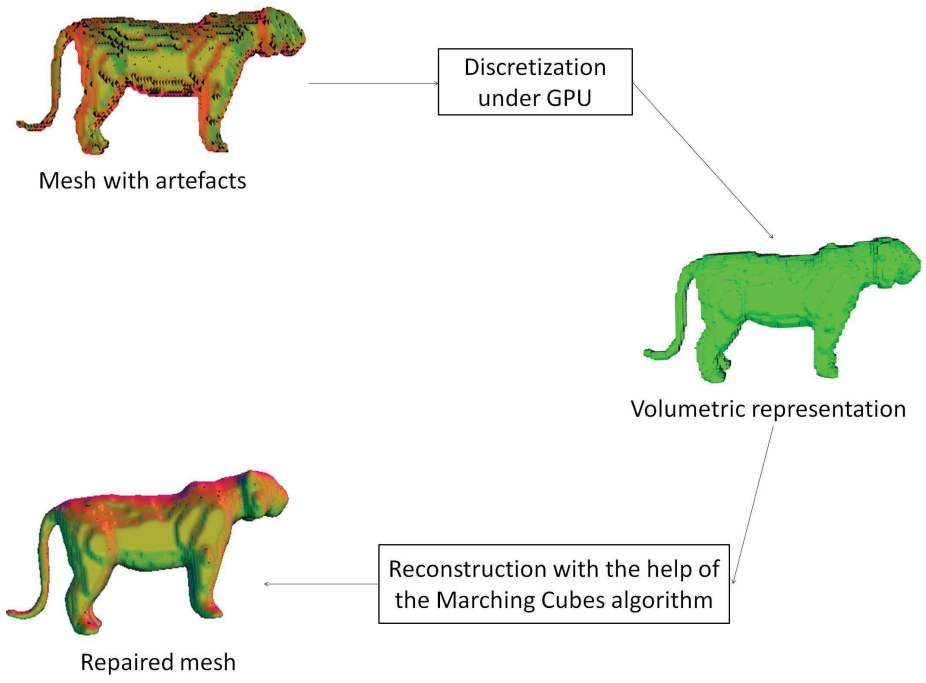

figure 1.7 – Process of mesh repairing.

However the 3D mesh representing the 3D object obtained after the repairing process, can have a high vertices resolution, which requires a large storage capacity in memory, and consequently, it makes the further processing more time-consuming. To overcome these shortcomings, a mesh vertices' decimation process must be performed in order to reduce the number of vertices, and to attribute the same number of vertices to the 3D objects in the collection. For this aim, we chose to work with the approach of Garland et al. [22]. In fact, it consists in decimating the mesh's vertices by using a quadratic error measurement which determines the cost of contracting a pair of vertices that belong to the mesh. For that purpose, this decimation process is performed in five steps. The first step consists in computing the quadratic error for each mesh vertex. In the second step, each two vertices that belong to an edge are selected, and considered as a valid pair for the contraction. The cost of contracting each valid pair is computed during the third step. In the fourth step, the valid pairs are inserted and sorted in a data structure according to the minimum cost. During the fifth step, the decimation process is performed by contracting iteratively the valid pair with a minimum cost.

1.4.2 Normalization of the 3D objects

The aim of this stage is to endow the features to be extracted invariance to articulations, to torsions, and to affine transforms. In fact, a 3D object can have an infinite number of articulations and torsions. Figure 1.8 shows examples of four possible articulations that a horse can have. If the articulations and torsions of a 3D

figure 1.8 – A horse with different articulations and torsions (3D models taken from Robert Sumner's collection of 3D shapes [44]).

object change, then the inter-vertices Euclidean distance changes. However, the displacement from a vertex v_i to another vertex v_j along the 3D object's mesh remains

almost the same, and it can be expressed by the geodesic distance between the two vertices v_i and v_j. For that purpose, the aim is to find an embedding of the mesh's vertices that best preserves the inter-vertices Euclidean distance. In other words, we have to find a space W, where for each w_i, $w_j \in W$ we have :

$$d_{geodesic}(v_i, v_j) = d_{Euclidean}(w_i, w_j) \tag{1.10}$$

Where w_i and w_j are respectively the embedding of the vertices v_i and v_j in the Euclidean space W. PCA preserves the covariance of the data points [58], and the multidimensional scaling finds an embedding that preserves the Euclidean distances between data points [24]. However, these data points can have a non-linear structure that cannot be visible by the PCA and the multidimensional scaling. For this reason, the kernel PCA [58, 65] appeared as an alternative of PCA in order to capture the hidden non-linear structure of the data points. But, its disadvantage is the choice of the kernel to use, because it depends on the non-linear structure to be identified from the data points [65]. In order to overcome the shortcomings of the previously mentioned methods, the ISOMAP embedding [14] is a non-linear dimensionality reduction method, which takes into consideration the non-linearity between the data points to preserve the Euclidean distance between them. To do so, the ISOMAP relies on the neighbourhoods of the data points and uses the geodesic distances between all its pairs. In the case of a $3D$ object, the targeted space is a set of its mesh's vertices linked by edges which constitute its connectivity graph. In fact, the ISOMAP proceeds in three steps. The first step consists in constructing the connectivity graph of the input mesh by finding the adjacent vertices of each vertex v_i. The second step consists in estimating the geodesic distance between the pairs of mesh's vertices. To do so, for each vertices' pair, the ISOMAP finds the shortest path along the mesh's graph which connects the two vertices with the help of Dijkstra's algorithm [25]. The estimated geodesic distance is computed from this shortest path, and it is stored in a matrix D_G of dimension $n \times n$, where n is the number of vertices, and $D_{G_{i,j}}$ denotes the geodesic distance between the vertex v_i and the vertex v_j. In the third step, the ISOMAP applies a multidimensional scaling to the matrix D_G. The aim is to construct an embedding of the mesh in a three-dimensional Euclidean space W that

best preserves the mesh's inter-vertices Euclidean distances. The vertices of the space W are chosen to minimize the following cost function :

$$E = \|\tau(D_G) - \tau(D_W)\|_{L^2} \tag{1.11}$$

Where D_W denotes the matrix of Euclidean distances $\{d_W(i,j) = \|w_i - w_j\|\}$, $\|M\|_{L^2}$ is the L^2 matrix norm, and τ is an operator that converts distances to inner products by the following formula :

$$\tau(D) = -\frac{1}{2}HSH \tag{1.12}$$

Where S is the matrix of squared distances $\{S_{i,j} = D^2_{i,j}\}$, and H is a centering matrix $\{H_{i,j} = \delta_{i,j} - 1/n\}$ with δ is an identity matrix of dimension $n \times n$. The role of the matrix H is to decrease the influence of any constant added in the elements of the distance matrix D [65]. The cost function E reaches its minimum when the coordinates of $w_i \in W$ are set to the top three eigenvectors of the matrix $\tau(D_G)$, where w_i represents the embedding of the vertex v_i in the space W. Figure 1.9 illustrates examples of obtaining almost the same articulations and torsions after the embedding with ISOMAP of three horses of the Robert Sumner's collection of $3D$ objects [44].

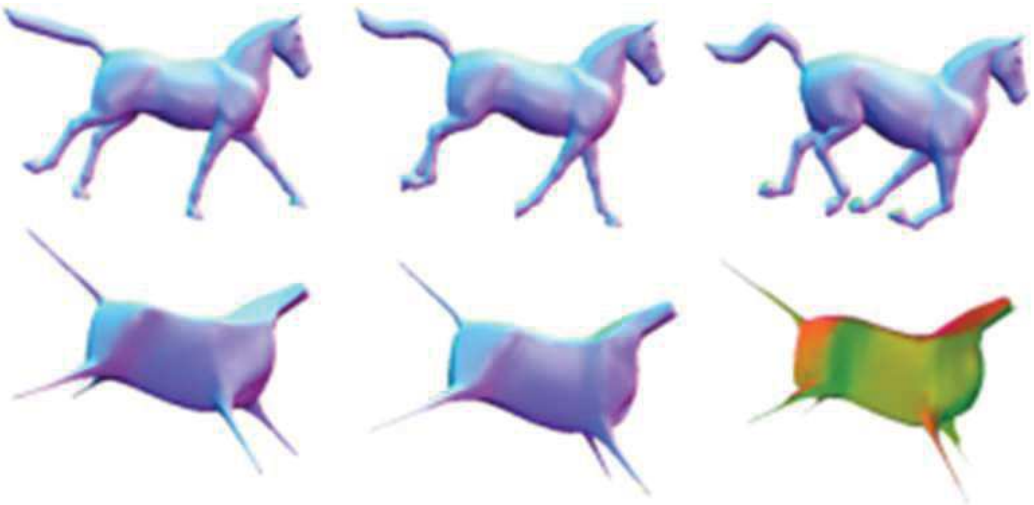

figure 1.9 – (Top) Three horses with different articulations and torsions, (Bottom) and their associated embedding with the ISOMAP.

However, the $3D$ objects can still have arbitrary positions, orientations and scales

in $3D$ space, even if they were embedded by the ISOMAP. On the one hand, the ISOMAP embedding does not have an influence on the original scale of the $3D$ object. Indeed, if a $3D$ object is put into a scale $\alpha \neq 1$, then $\{v'_k = \alpha v_k\}_{k=1,\ldots,n}$, and consequently the geodesic distance between each vertex pair v'_i and v'_j is :

$$d_{geodesic}(v'_i, v'_j) = \alpha d_{geodesic}(v_i, v_j) \neq d_{Euclidean}(w_i, w_j) \tag{1.13}$$

On the other hand, for a given $3D$ object with different articulations and torsions, the ISOMAP embedding's role is to find a vertices' space W, where the inter-vertices Euclidean distance is preserved. But, this does not ensure associating the same orientation for $3D$ objects which belong to the same category. Because the principal axis of the resulting space W do not coincide with those of the benchmark reference [14, 65].

In order to capture the invariant features to affine transforms for each $3D$ object, a feasible scheme is to center, to scale, and to orient the $3D$ model by using the PCA [15]. To do so, the PCA proceeds in three steps. The first step consists in centering the $3D$ objects in order to have the same position in the $3D$ space, and to endow the features invariance to translation. The centering is performed through the following formula :

$$P'_i = P_i - G = (x_i - x_G, y_i - y_G, z_i - z_G) \tag{1.14}$$

Where P_i is a vertex of the $3D$ object, and G is the mass center of the $3D$ object defined by the following formula :

$$G = \frac{\sum_{i=1}^{N} s_i g_i}{\sum_{i=1}^{N} s_i} \tag{1.15}$$

Where N is the number of triangles, and s_i and $g_i = (g_{ix}, g_{iy}, g_{iz})$ are respectively the surface and center of the i^{th} triangle of the mesh that represents the $3D$ object.

In the second step, the $3D$ object is scaled in order to be bound by a unit sphere. The principle is to search the farthest vertex from the benchmark's center O and to compute the associated distance by the following formula :

$$d_m = \max_{i=1,...,n} d(O, P'_i) \tag{1.16}$$

Where $d(O, P'_i)$ is the Euclidean distance between the vertex P'_i and the benchmark's center O, and d_m is the distance of the farthest vertex to O. The scaling is performed through the following formula :

$$P''_i = \frac{1}{d_m} P'_i = \left(\frac{x'_i}{d_m}, \frac{y'_i}{d_m}, \frac{y'_i}{d_m} \right) \tag{1.17}$$

The third step consists in aligning the $3D$ objects in order to have the same orientation. For that purpose, the mesh covariance analysis is performed to determine the principal axes of the $3D$ object. To do so, the mesh's covariance matrix is computed through the following formula :

$$C = \frac{1}{N} \sum_{i=1}^{N} S_i g_i g_i^T = \frac{1}{N} \sum_{i=1}^{N} S_i \begin{pmatrix} g_{ix}^2 & g_{ix}g_{iy} & g_{ix}g_{iz} \\ g_{iy}g_{ix} & g_{iy}^2 & g_{iy}g_{iz} \\ g_{iz}g_{ix} & g_{iz}g_{iy} & g_{iz}^2 \end{pmatrix} \tag{1.18}$$

The covariance matrix C is symmetric, consequently its eigenvalues are positives. The eigenvalues of the matrix C and their associated eigenvectors are sorted in descending order. These eigenvectors represent the principal axes of the $3D$ object and they are normalized to form the rows of the alignment matrix R. The alignment of the $3D$ object is performed as follows :

$$P''' = P''R = \{P''_i R, i = 1, ..., n\} \tag{1.19}$$

The different normalization steps of the $3D$ object can be summarized in Figure 1.10.

1.5 Experimental results

At first, the proposed method was experimentally evaluated using the manifold version of the $3D$ object collection of McGill [19]. This latter is composed of 457 $3D$ models in high-level semantics 19 categories, where the $3D$ objects of the same cate-

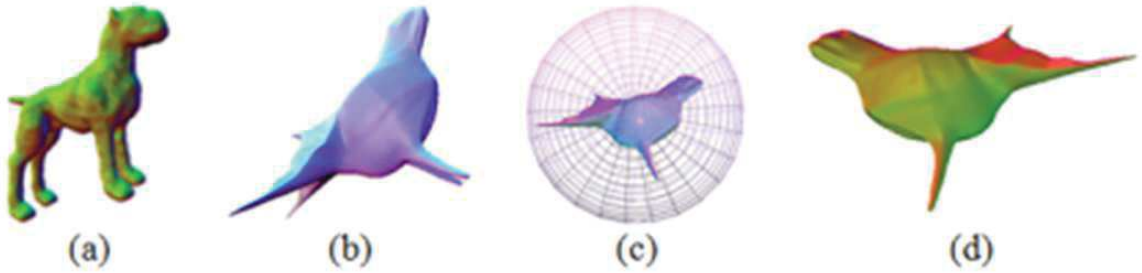

figure 1.10 – (a) An input 3D object, (b) the representation of the 3D object in the space W after its embedding by the ISOMAP, (c) centering and scaling inside the unit sphere, (d) 3D alignment using PCA.

gory have different shapes, articulations, torsions, positions, scales, and orientations. Figure 1.11 illustrates examples of 3D objects of each category of the 19 classes of the 3D objects collection of McGill. In fact, it is one of the standard 3D models collection used to validate several descriptors of 3D objects [1, 3, 11, 12].

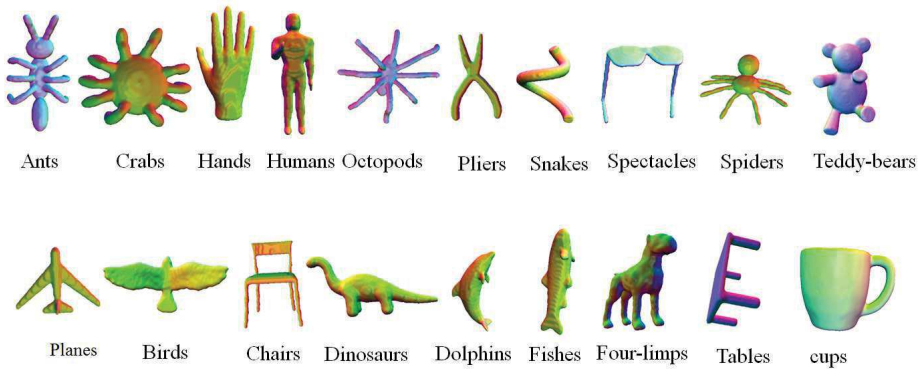

figure 1.11 – Examples of 3D models from the 3D object collection of McGill.

Afterwards, the proposed method was experimentally evaluated by merging different collections of 3D objects. For that purpose, we extended the 3D object collection of McGill. by 3D models from the 3D shape benchmark of architectural data [64], by the articulated 3D objects from Robert Sumner's collection of 3D shapes [44], and by 3D models from the Princeton's collection of 3D objects [62]. The obtained collection of 3D objects has 2902 models and 121 categories. Our 3D object recognition system can be summarized in Figure 1.12. At first, for each 3D object in the collection

of $3D$ models, a preprocessing stage is performed to repair its mesh and to endow the features invariance to affine transforms, articulations, and torsions. Afterwards, the $3D$ object is described by a set of 18 size functions which will be stored in the feature dataset. Thus, the comparison between the $3D$ objects is performed through the similarity measure between their associated sets of 18 size functions. Finally, a soft decision process is performed in order to retrieve the $3D$ models according to their resemblance to the $3D$ query object.

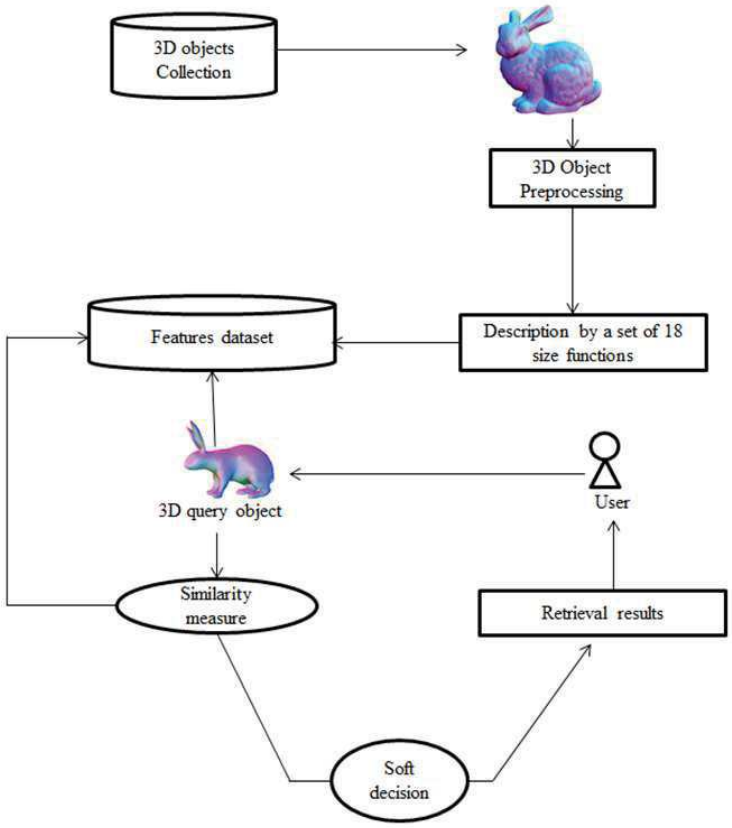

figure 1.12 – The $3D$ object recognition system.

To evaluate the proposed method, each $3D$ model is used as a query object, and a soft decision process is performed to sort the $3D$ objects of the collection according to their resemblance to the $3D$ model to be recognized. The recognition performance was evaluated in terms of recall and precision [26]. For that purpose, a query is formed by a $3D$ object belonging to a category C_i of Q $3D$ models. The recall represents

the number of $3D$ objects belonging to the category C_i from the k first $3D$ objects retrieved, we denote it by $N_{corrects}(k)$, on the number Q of the $3D$ objects of the category C_i. The precision is the number of $3D$ objects which belong to the category C_i found from the k first retrieved $3D$ objects, on the number of responses k. The precision and recall can be formulated as follows :

$$Precision(k) = \frac{N_{corrects}(k)}{k}, \tag{1.20}$$

$$Recall(k) = \frac{N_{corrects}(k)}{Q} \tag{1.21}$$

We compared the proposed method with its global version which relies on the description of the $3D$ object by one size function resulting from the measuring function $\varphi(P_i) = e^{-d_{Euclidian}(O,P_i)}$, with the work of Biasotti et al.[12] where they described a $3D$ object by a multidimensional size function, the work of Ion et al.[1] where the $3D$ object is described by an eccentricity histogram, and the local approach of Xiaofeng et al.[71] where the $3D$ object is described by the local areas of its vertices. We implemented the mentioned methods in C++ programming language.

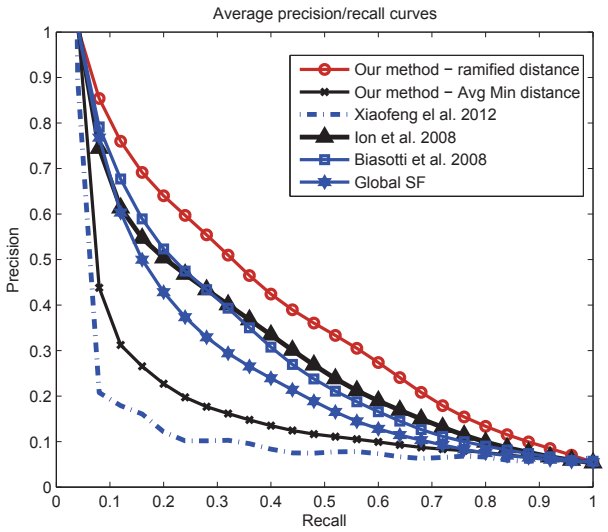

figure 1.13 – The averages precisions/recalls curves for the 457 $3D$ models of the $3D$ object collection of McGill.

tableau 1.2 – The average recognition scores of different methods for the 457 3D models of the 3D object collection of McGill.

Methods	First tier	Second tier
Our method	40.88%	57.17%
Zimmer et al. 2013	26.6%	40.0%
Lian et al. 2012	24.9%	38.5%
Xiaofeng et al. 2012	14.98%	22.81%
Ion et al. 2008	35.70%	50.37%
Biasotti et al. 2008	35.42%	48.24%

Figure 1.13 and Figure 1.14 represent, respectively the averages precision/recalls curves of our method, the work of Biasotti et al. [12], the work of Ion et al. [28], and the local approach of Xiaofeng et al. [71] for the 457 3D models of the 3D object collection of McGill, and for the 2902 3D objects of the extended collection. Table 1.2 summarizes the accuracy of our method, the scores obtained by the method Zimmer et al. [80], the scores obtained by the method of Lian et al. [79], and the other approaches for the 457 3D models of the 3D object collection of McGill by the means of the first tier and the second tier evaluation measures [19]. In fact, the first tier has the same principle of the recall, because it provides the recognition score, where the number of retrieved 3D objects is the number of 3D objects in the category of the 3D query model. The second tier provides the recognition score, where the number of retrieved 3D objects is twice the number of 3D objects in the category of the 3D query model. As illustrated in Figure 1.13 and Table 1.2, our method got the highest scores when it was compared with the other approaches of the related work for the 457 3D models of the 3D object collection of McGill. From Figure 1.14, the scores obtained by our method are favourably comparable to the related work when the recall is upper than 0.1 when it was validated on the extended 3D objects collection. This illustrates the relevant local spatial information provided by the critical points of each portion of the 3D object.

Table 1.3 summarizes the scores gotten by our method the mean of first tier, the different combinations of the subsets of portions according to the benchmark axis, the global version of our method, and the three methods of the related work for the

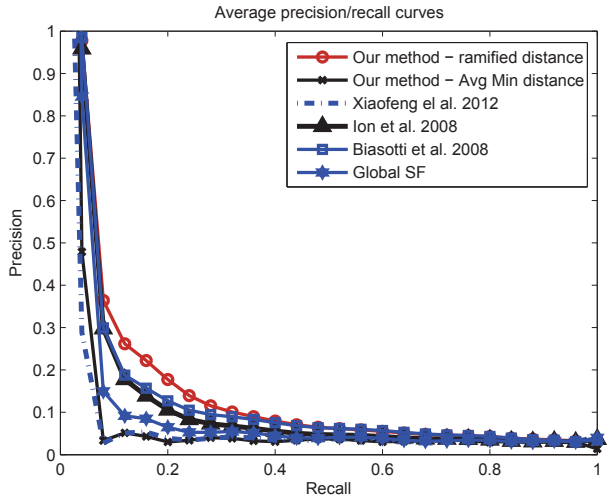

figure 1.14 – The averages precisions/recalls curves for the 2902 $3D$ models of the extended collection of $3D$ objects.

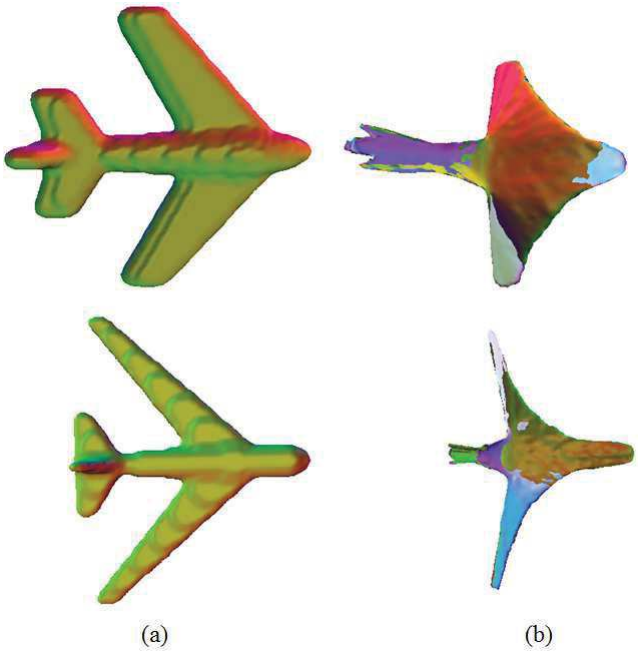

figure 1.15 – (a) Two airplanes with significant difference in their shapes, (b) different articulations after the ISOMAP embedding.

tableau 1.3 – The recognition scores of different methods for the 19 categories of the $3D$ object collection of McGill.

Classes	S_X	S_Y	S_Z	S_X, S_Y, and S_Z	S	Our method with average minimum distance	Our method with ramified distance	Global version of our method	Ion et al.	Xiaofeng et al.	Biasotti et al.
Ants	31.22%	27.33%	22.56%	38.00%	30.67%	3.78%	32.33%	34.78%	**43.33%**	8.33%	20.44%
Crabs	39.89%	37.67%	44.78%	39.56%	53.44%	24.22%	59.44%	44.78%	**60.22%**	43.11%	31.67%
Hands	16.00%	22.00%	21.00%	18.25%	24.75%	6.00%	**25.50%**	17.50%	24.00%	11.75%	13.00%
Humans	33.41%	32.82%	31.27%	31.87%	36.50%	5.35%	**44.35%**	9.87%	35.79%	5.95%	23.78%
Octopuses	13.28%	17.76%	18.40%	14.88%	17.44%	5.60%	22.08%	23.36%	18.88%	14.08%	**38.56%**
Pliers	55.75%	29.75%	40.25%	42.75%	51.25%	37.50%	62.75%	**67.50%**	66.75%	13.75%	65.75%
Snakes	47.84%	41.44%	48.16%	50.24%	57.12%	55.68%	59.20%	34.88%	**77.60%**	27.20%	27.36%
Spectacles	46.08%	44.96%	85.28%	66.24%	68.64%	65.44%	57.12%	59.36%	**74.72%**	18.88%	61.12%
Spiders	29.97%	31.84%	37.67%	47.35%	39.23%	18.00%	37.57%	40.69%	35.07%	5.93%	**58.79%**
Teddy-bears	50.00%	31.50%	35.75%	**56.00%**	52.00%	8.75%	54.00%	29.75%	24.50%	20.25%	30.50%
Airplanes	21.01%	14.05%	15.53%	17.31%	16.42%	6.07%	15.68%	13.02%	12.13%	6.07%	**26.04%**
Birds	**46.71%**	20.86%	21.09%	29.93%	36.73%	12.02%	42.86%	33.11%	17.69%	14.74%	26.53%
Chairs	34.03%	20.98%	33.65%	43.86%	48.39%	19.47%	41.97%	22.68%	28.54%	15.50%	**54.44%**
Cups	20.96%	20.80%	22.72%	21.76%	28.32%	24.96%	37.28%	25.44%	**49.92%**	27.84%	39.68%
Dinosaurs	30.19%	18.01%	27.15%	30.47%	37.67%	11.91%	**42.11%**	19.94%	30.75%	15.24%	34.90%
Dolphins	27.78%	22.92%	27.08%	25.69%	31.25%	20.83%	**38.19%**	26.39%	11.81%	9.72%	27.78%
Fishes	31.00%	32.51%	16.26%	34.59%	39.32%	44.42%	**55.39%**	33.65%	41.78%	5.48%	44.99%
Four-limbs	23.41%	21.12%	16.86%	**25.81%**	22.06%	23.62%	22.16%	17.38%	13.84%	6.97%	21.44%
Tables	15.29%	17.36%	9.09%	11.78%	18.39%	16.94%	**26.65%**	24.79%	10.95%	13.84%	26.24%
Average recognition scores	32.31%	26.61%	30.24%	34.02%	37.35%	21.61%	**40.88%**	30.47%	35.70%	14.98%	35.42%

19 categories of the $3D$ objects collection of McGill. In Table 1.3, our method where the $3D$ object description relies on the set of 18 size functions and the similarity measure is based on the ramified distance that got the best average score. However, some categories such as spectacles and chairs need respectively four size functions associated to subset S_Z, and six size function associated to the subset S in order to get a high score. Besides, the usage of the average minimum distances of our method decreased the average recognition score, because the subsets S, S_X, S_Y, and S_Z do not have the same relevance. Our method got the lowest score in the airplanes' category, because the $3D$ objects that belong to it have a significant difference between their shapes. Indeed, in the normalization stage, the ISOMAP requires the computation of geodesic distances of the vertex pairs by finding the shortest path between them. Where the $3D$ objects have significant differences between their shapes, their ISOMAP embeddings outcome different articulations and torsions as it is illustrated in Figure 1.15-(b). As a result, the $3D$ objects of the airplanes' category got different portions

after the splitting process, and consequently, different descriptions by the set of 18 size functions.

tableau 1.4 – The average recognition scores of different methods for the 2902 3D objects of the extended collection.

Methods	First tier	Second tier
Our method with ramified distance	13.24%	18.18%
Ion et al. 2008	11.11%	14.59%
Biasotti et al. 2008	10.77%	14.35%
Global version of our method	8.57%	11.11%
Our method with average minimum distance	6.14%	7.50%
Xiaofeng et al. 2012	5.36%	6.83%
Our method with only the subset S_X	8.00%	10.75%
Our method with only the subset S_Y	6.27%	7.72%
Our method with only the subset S_Z	6.21%	7.69%
Our method with the subsets S_X, S_Y, and S_Z	7.02%	9.06%
Our method with only the subset S	7.47%	9.84%

Table 1.4 summarizes the accuracy of our method, and the other approaches for the 2902 3D objects of the extended collection by the means of the first tier, and the second tier evaluation measures [62]. On the Table 1.4, our method that relies on ramified distance got the highest recognition scores on all the other approaches. Meanwhile, using the average minimum distance on our method could not get the best scores, because as in the case of the 3D objects collection of McGill, the subsets S, S_X, S_Y, and S_Z do not have the same relevance for all the 3D objects in the extended collection. Besides, the other combinations of the sets of portions according to the benchmark axis could not reach the highest recognition scores comparing to the other approaches of the related work. Consequently, the use of the combination of 18 portions with the ramified distance was necessary to improve the recognition score of our method.

Figure 1.16 represents the average computation time of the different methods obtained on a PC with 2.80 Ghz Intel Xeon® CPU with RAM 12.2 Go and NVIDIA Tesla c2050/c2070 GPU computing processors, and running under Windows 7 64 bits. According to Figure 1.16, our method is more computation time-consuming than the other approaches, because it requires the normalization stage and the splitting of the 3D object into 18 portions in order to represent each one of them by a size function.

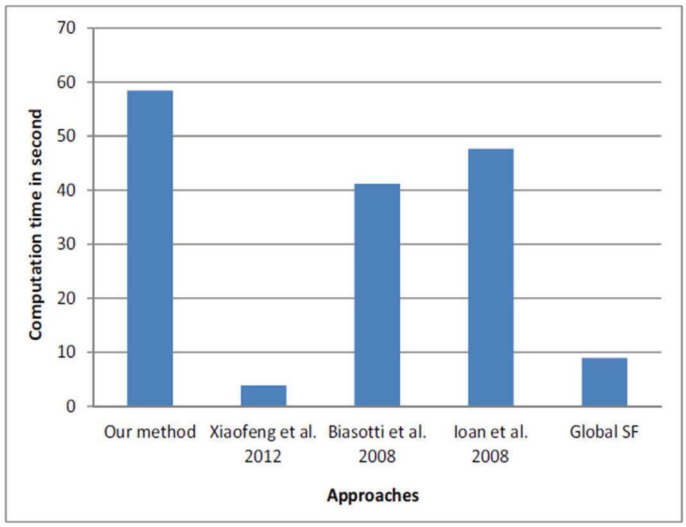

figure 1.16 – Computational time of different methods.

1.6 Conclusions

In this paper, we developed a new local approach which describes the $3D$ object by a set of 18 size functions. The main idea of our method is the local usage of the topological invariants provided by the critical points to describe the $3D$ object. For that purpose, the $3D$ object is split into 18 portions, to represent each one of them by a size function resulting from a measure function, which captures the concavities and the convexities features of the portion. At first, the method was validated by using the $3D$ object collection of McGill of 457 models. The obtained scores outperform the related work. Afterwards, the method was validated by using an extended $3D$ objects collection of 2902 models. The obtained scores are favourably comparable to the related work. However, our method is more computation time-consuming than the other approaches, because it requires a normalization stage, the splitting of the $3D$ object into 18 portions, and the description of each portion by a size function splitting the targeted $3D$ object into 18. To overcome this limitation, we are planning to parallelize the size function's computation on GPU in order to accelerate the computation time of our method.

Bibliographie

[1] A. Ion, N.M. Artner, G. Peyré, S.B.L. Marmol, W.G. Kropatsch, L.D. Cohen. 3D Shape Matching by Geodesic Eccentricity. Proceeding of Computer Vision and Pattern Recognition Workshops, pp. 1-8, 2008.

[2] J. Assfalg, A. Del Bimbo, P. Pala. Content-based Retrieval of 3D Models through Curvature Maps : A CBR Approach Exploiting Media Conversion. Multimedia Tools and Applications, 31(1), pp. 29-50, 2006.

[3] J. Rabin, G. Peyré, L.D. Cohen. Geodesic Shape Retrieval via Optimal Mass Transport. Proceeding of the 11th European Conference on Computer Vision, Springer-Verlag, pp. 771-784, 2010.

[4] M. Ferri, S. Lombardini, C. Pallotti. Leukocyte classification by size functions. In Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, IEEE Computer Society Press, pp. 223-229, 1994.

[5] W.E. Lorensen, H.E. Cline. Marching Cubes : A high resolution 3D surface construction algorithm. Computer Graphics, 21(4), 1987.

[6] D. V. Vranic, D. Saupe. 3D Shape Descriptor Based on 3D Fourier Transform. EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS '01), pp. 271-274, 2001.

[7] D. Vranic, 3D Model Retrieval, Ph.D. dissertation, University of Leipzig, 2004.

[8] P. Frosini, C. Landi. Size Theory as a Topological Tool for Computer Vision. Pattern Recognition and Image Analisys, 9(4), pp. 596-603, 1999.

[9] M. Handouyahia, D. Ziou, S. Wang. Sign language recognition using moment-based size functions. In : Proc. of Vision Interface 99, pp. 210-216, 1999.

[10] A. Verri, C. Uras. Metric-topological approach to shape representation and recognition. Image and Vision Computing. 14, pp.189–207, 1996.

[11] S. Biasotti, D. Giorgi, M. Spagnuolo, B. Falcidieno. Size functions for comparing 3D models. Pattern Recognition, Elsevier, 41(9), pp. 2855-2873, 2008.

[12] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, C. Landi. Multidimensional Size Functions for Shape Comparison. Journal of Mathematical Imaging and Vision, Springer, 32(2), pp. 161 - 179, 2008.

[13] M. D'Amico, P. Frosini, C. Landi, Natural pseudo-distance and optimal matching between reduced size functions, Acta Applicandae Mathematicae, 109(2), pp. 527-554, 2010.

[14] J.B. Tenenbaum, V. de Silva, J. Langford. A global geometric framework for nonlinear dimensionality reduction. Science 290 : 22, pp. 2319-2323, 2000.

[15] D.Y. Chen, M. Ouhyoung. A 3D model alignment and retrieval system. Proceeding of International Computer Symposium, 2002.

[16] S. Biasotti, L. De Floriani, B. Falcidieno, P. Frosini, D. Giorgi, C. Landi, L. Papaleo, M. Spagnuolo. Describing shapes by geometrical-topological properties of real functions. Computing Surveys, ACM, 40(4), 2008.

[17] R.C. Veltkamp. Shape Matching : Similarity Measure and Algorithms. Proceedings Shape Modelling International, IEEE Press, pp. 188-197, 2001.

[18] B.J. Super. Learning Chance Probability Functions for Shape Retrieval or Classification. Proceedings of the IEEE Workshop on Learning in Computer Vision and Pattern Recognition, 2004.

[19] McGill 3D Shape Benchmark http ://www.cim.mcgill.ca/ shape/benchMark/

[20] J.W. H. Tangelder, R.C. Veltkamp. A survey of content based 3D shape retrieval methods. Multimedia Tools and Applications, 39, pp. 441-471, 2008.

[21] B. Bustos, D. Keim, D. Saupe, T. Schreck. Content-based 3D object retrieval. IEEE Computer Graphics and Applications, IEEE Computer Society, 27(4), pp. 22-27, 2007.

[22] M. Garland, P.S. Heckbert. Surface simplification using quadric error metrics. The 24th Annual Conference on Computer Graphics and interactive Techniques, pp. 209-216, 1997.

[23] M. D'Amico. A new optimal algorithm for computing size function of shapes. In CVPRIP Algorithms III : Proceedings of the International Conference on Computer Vision, Pattern Recognition and Image Processing, pp. 107-110, 2000.

[24] I. Borg, P. Groenen. Modern Multidimensional Scaling : Theory and Applications. Springer Series in Statistics, 1996.

[25] E.W. Dijkstra. A note on two problems in connexion with graphs. Numerische Mathematik, 1, pp. 269-271, 1959.

[26] D.L. Olson, D. Delen. Advanced Data Mining Techniques. Springer, 1st edition, 2008.

[27] T. Akenine-Möller. Fast 3D triangle-box overlap testing. In the proceeding of ACM SIGGRAPH Courses, 2005.

[28] M. Botsch, M. Pauly, C. Rössl, S. Bischoff, L. Kobbelt. Geometric Modeling Based on Triangle Meshes. SIGGRAPH Course Notes. 2006.

[29] J. Ponce, M. Hebert, C. Schmid, A. Zisserman (Eds.). Toward Category-Level Object Recognition. Lecture Notes in Computer Science 4170, ISBN 3-540-68794-7. Springer 2006.

[30] S. Dickinson, A. Leonardis, B. Schiele, M. Tarr, (Editors). Object Categorization : Computer and Human Vision Perspectives. Cambridge University Press, 2009.

[31] G. Flitton, T. Breckon, N.M. Bouallagu. Object Recognition using 3D SIFT in Complex CT Volumes. Proceedings of the British Machine Vision Conference, pp. 11.1-11.12, BMVA Press, 2010.

[32] U.Bagci, X. Chen, J.K.Udupa. Hierarchical Scale-Based Multi-Object Recognition of 3D Anatomical Structures. IEEE Transactions on Medical Imaging, 31(3), 2012.

[33] Y. Takahashi, Y. Muramatu, K. Kato. Object Recognition using Simultaneous Pattern-Matching in Virtual Environment. SICE-ICASE International Joint Conference, pp. 5579-5583, 2006.

[34] G. Rajamanoharan, S. Zafeiriou, M. Pantic, L. Yin. Static and Dynamic 3D Facial Expression Recognition : A Comprehensive Survey. Image and Vision Computing. 30(10), pp. 683-697, 2012.

[35] A. Ruifrok, A. Scheenstra, R.C. Veltkamp. A Survey of 3D Face Recognition Methods. Audio and Video-based Biometric Person Authentication. Lecture Notes in Computer Science, 3546, pp. 325-345, 2005.

[36] F. Stanco, S. Battiato. Digital Imaging for Cultural Heritage Preservation : Analysis, Restoration, and Reconstruction of Ancient Artworks. CRC Press, 2011.

[37] F. Bosché. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. Advanced Engineering Informatics, 24(1), 107-118, 2010.

[38] E. Valero, A. Adan, C. Cerrada. Automatic Construction of 3D Basic-Semantic Models of Inhabited Interiors Using Laser Scanners and RFID Sensors. Sensors, 12(5), pp. 5705-5724, 2012.

[39] G. Sansoni, M. Trebeschi, F. Docchio. State-of-The-Art and Applications of 3D Imaging Sensors in Industry, Cultural Heritage, Medicine, and Criminal Investigation. Sensors, 9(1), pp. 568-601, 2009.

[40] E. Pérez, S. Salamanca, P. Merchán, A. Adán, C. Cerrada, I. Cambero. A Robust Method for Filling Holes in 3D Meshes Based on Image Restoration. Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science, 5259, pp. 742-751, 2008.

[41] G. Barequet, M. Sharir. Filling Gaps in the Boundary of a Polyhedron. Computer Aided Geometric Design, 12(2), pp. 207-29, 1995.

[42] P. Borodin, M. Novotni, R. Klein. Progressive Gap Closing for Mesh Repairing. Advances in Modelling, Animation and Rendering, edited by J. Vince and R. Earnshaw, pp. 201-13. Springer Verlag, 2002.

[43] B. Schölkopf, A.J. Smola, K.R. Müller. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. Neural Computation, Volume 10, Issue 5, pp. 1299-1319, 1998.

[44] Robert Sumner's shapes repository http ://people.csail.mit.edu/ sumner/research/deftransfer/

[45] M. D'Amico, M. Ferri, I. Stanganelli. Qualitative Asymmetry Measure for Melanoma Detection. IEEE International Symposium on Biomedical Imaging , pp. 1155-1158, 2004.

[46] M. Dutton, R. Doran. Autodesk 3ds Max 8 Revealed. Course Technology Inc, 2005.

[47] D. Abbott. AutoCAD : Secrets Every User Should Know. Sybex ; 1 edition, 2006.

[48] R. Baeza-Yates, G. Valiente. An Image Similarity Measure Based on Graph Matching. Proceedings of the Seventh International Symposium on String Processing Information Retrieval, pp. 28- 38, 2000.

[49] A. Pinz. Object Categorization, Foundations and Trends® in Computer Graphics and Vision, 1(4), Pages : 255-353, 2006.

[50] P. Zezula, G. Amato, V. Dohnal, M. Batko. Similarity Search : The Metric Space Approach. The Kluwer International Series on Advances in Database Systems, Springer, Volume 32, 2006.

[51] J. S. Beis and D. G. Lowe. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1000-1006, 1997.

[52] J. S. Beis and D. G. Lowe. Indexing without invariants in 3D object recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 21, Issue 10, pp. 1000-1015, 1999.

[53] T. Tung and F. Schmitt. The augmented multi-resolution Reeb graph approach for content based retrieval of 3D shapes. International Journal of Shape Modeling (IJSM), Volume 11, Issue 1, pp. 91-120, 2005.

[54] M. Hilaga, Y. Shinagawa, T. Kohmura, T.L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In SIGGRAPH'01 : Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. ACM Press, pp. 203-212, 2001.

[55] S. Obdrzalek, J. Matas. Sub-linear Indexing for Large Scale Object Recognition. Proceedings of the British Machine Vision Conference BMVC, pp. 3.1–3.10, 2005.

[56] S. Mahmoudi, M. Daoudi. 3D Models Retrieval by using Characteristic Views. The International conference on Pattern Recognition, pp : 457-460, 2002.

[57] F. Mokhtarian, K. Mackworth. A Theory of Multiscale, Curvature-Based Shap Representation for Planar Curves. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(8), pp. 789-805, 1992.

[58] S. Theodoridis, K. Koutroumbas. Pattern Recognition, 4th Edition. Academic Press, 2008.

[59] H. Laga, H. Takahashi, M. Nakajima. Spherical Wavelet Descriptors for Content-based 3D Model Retrieval. IEEE International Conference on Shape Modeling and Applications (SMI2006), pp. 75-85, 2006

[60] A. Arampatzis, J. Kamps, S. Robertson. Where to stop reading a ranked list ? : threshold optimization using truncated score distributions. Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 524-531, 2009.

[61] C. Villani. Topics in Optimal Transportation. American Mathematical Society, 2003.

[62] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser. The Princeton Shape Benchmark. In the proceedings of Shape Modeling International, pp. 167-178, 2004.

[63] M. Attene. A lightweight approach to repairing digitized polygon meshes. International Journal of Computer Graphics, 26(11), pp. 1393-1406, 2010.

[64] R. Wessel, I. Blümel, R. Klein. A 3D Shape Benchmark for Retrieval and Automatic Classification of Architectural Data. In proceedings of Eurographics 2009, Workshop on 3D Object Retrieval, pp. 53-56, 2009.

[65] N. Zheng, J. Xue. Statistical Learning and Pattern Analysis for Image and Video Processing. Springer, 2009.

[66] D.H. Eberly. Geometric Methods for analysis of Ridges in N-dimensional Images. Ph.D. thesis, University of North Carolina at Chapell Hill, 1994.

[67] S. Suresh, N. Sundararajan, P. Saratchandran. Risk-sensitive loss functions for sparse multi-category classification problems. Information Sciences, 178 (12), pp. 2621-2638, 2008.

[68] J.J. Freeman. Experiments in discrimination and classification. Pattern Recognition, 1(3), pp. 207-218, 1969.

[69] P. Frosini. Measuring shapes by size functions. In Intelligent Robots and Computer Vision X : Algorithms and Techniques, 1607, pp. 122-133, 1991.

[70] A. Verri, C. Uras, P. Frosini, M. Ferri. On the use of size functions for shape analysis. Biolog. Cybernet. 70, pp. 99-107, 1993.

[71] C. Xiaofeng, G. Chuanwei, Z. Xutang, J. Lijun. 3D Model Retrieval Based on Projected Area at Mesh Vertex. Third International Conference on Digital Manufacturing and Automation, IEEE International. pp. 1-4, 2012.

[72] K. Mamou, F. Ghorbel. A simple and efficient approach for 3D mesh approximate convex decomposition. IEEE International Conference on Image Processing, pp. 3501-3504, 2009.

[73] R. Lium, H. Zhang. Mesh Segmentation via Spectral Embedding and Contour Analysis. Computer Graphics Forum, 26, pp. 385-394, 2007.

[74] X. Pan, Q.H. Chen, Z. Liu. 3D shape recursive decomposition by Poisson equation. Pattern Recognition Letters, 30(1), pp. 11-17, 2009.

[75] D. Gallup, J.-M. Frahm, M. Pollefeys. Piecewise Planar and Non-Planar Stereo for Urban Scene Reconstruction. IEEE Conference Computer Vision and Pattern Recognition, pp. 1418-1425, 2010.

[76] O.R. Terrades, E. Valveny, S. Tabbone. Optimal Classifier Fusion in a Non-Bayesian Probabilistic Framework.IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(9) ,pp. 1630-1644, 2009.

[77] P. Merrell, A. Akbarzadeh, L. Wang, P. Mordohai, J.-M. Frahm, R. Yang, D. Nistér, M. Pollefeys. Real-Time Visibility-Based Fusion of Depth Maps. IEEE 11th International Conference on Computer Vision (ICCV), pp. 1-8, 2007.

[78] P. Frosini, M. Pittore. New methods for reducing size graphs. International Journal of Computer Mathematics, 70(3), pp.505-517, 1999.

[79] Z. Lian, A. Godil, P. Rosin, X. Sun. A new convexity measurement for 3d meshes. IEEE CVPR, pp. 119-126, 2012. .

[80] H. Zimmer, M. Campen, L. Kobbelt. Efficient Computation of Shortest Path-Concavity for 3D Meshes. IEEE CVPR, pp. 2155-2162, 2013.

Bibliographie

Chapitre 2

La reconnaissance d'objets $3D$ basée sur une fonction de taille résultante d'une caractéristique d'invariants topologiques

Dans l'approche locale présentée dans le chapitre 1, chaque objet $3D$ est représenté par un ensemble de 18 fonctions de taille 1D. Cependant, elle présente quelques limitations. D'une part, le processus de découpage d'objets $3D$ est sensible, puisqu'il s'effectue selon les axes principaux de l'objet $3D$ obtenus par l'analyse en composante principale. En effet, si le même objet $3D$ manque d'une ou plusieurs portions, ses axes principaux vont changer, et par conséquent, les portions obtenues après le découpage seront différentes à celles obtenus dans le cas où l'objet $3D$ est complet. Par suite, sa description par un ensemble de 18 fonctions ne sera pas le même dans le cas où l'objet $3D$ ne manque d'aucune portion. Ainsi, le processus de description d'objets $3D$ par un ensemble de 18 fonctions de taille décrit dans le chapitre 1 exige le fait que les objets $3D$ soient complets, ce qui n'est pas toujours le cas. Pour cette raison, l'approche présentée dans le chapitre 1 ne peut pas être utilisée pour la reconnaissance partielle d'objets $3D$. D'une autre part, l'approche locale exige un processus de normalisation, puisque la fonction de mesure utilisée n'est pas invariante aux ar-

ticulations, aux torsions et aux transformations affines. Par conséquent, le processus de la description d'objets 3D de l'approche locale est coûteux en temps de calcul. Afin de surpasser ces limitations, dans ce chapitre nous présentons une méthode de reconnaissance d'objets 3D à travers une seule fonction de taille résultante d'une caractéristique qui incorpore l'invariance aux transformations affines, articulations et torsions. Pour cela, chaque objet 3D est décrit globalement par une seule fonction de taille résultante d'une fonction de mesure qui représente la caractéristique du déplacement surfacique. Par conséquent, la méthode de description n'aura pas besoin de processus de normalisation. En plus, la méthode peut être utilisée pour la reconnaissance d'objets 3D complets et incomplets. La validation de la méthode s'est effectuée en utilisant la collection d'objets 3D de McGill, la collection étendue de 2902 objets 3D et la collection d'objets 3D de Rodola et al. [19] pour la reconnaissance partielle d'objets 3D. Les scores obtenus sont supérieurs à ceux obtenus dans l'état de l'art.

J'ai développé, implanté, validé et rédigé la totalité de ce travail sous la supervision du Professeur Djemel Ziou. Dans les pages qui suivent, nous présentons notre méthode dans un article intitulé 3D **object recognition through a size function resulting from an invariant topological feature**. Cet article est en cours de révision dans la revue **IEEE Transactions on Pattern Analysis and Machine Intelligence**. La version manuscrite de cet article a été modifiée suite aux recommandations des membres de jury et cela n'engage que ma responsabilité.

$3D$ object recognition through a size function resulting from an invariant topological feature

Mohammed Ayoub Alaoui Mhamdi
Département d'informatique, Université de Sherbrooke,
Sherbrooke, Québec, Canada J1K 2R1
ayoub.alaoui@usherbrooke.ca

Djemel Ziou
Département d'informatique, Université de Sherbrooke,
Sherbrooke, Québec, Canada J1K 2R1
djemel.ziou@usherbrooke.ca

Keywords: $3D$ model description, object recognition, object categorization, shape classification.

Abstract

In this paper, a critical points based descriptor for $3D$ objects recognition is presented. It is based on the topological invariants provided by the critical points of the $3D$ object. The critical points and the links between them are represented by a size function resulting from a measure function that captures the surface displacement along the $3D$ object, and that encompasses invariance to affine transformations, articulations and torsions. In order to tackle the problems of partial matching of the $3D$ objects, a well-suited metric learning method is used to weight the matchings according to their relevance. The proposed method's performance was validated by different collections of $3D$ objects. The obtained scores are favourably comparable to the related work.

2.1 Introduction

Since the last decade [29, 30], the $3D$ object recognition has gained more importance on various applications such as medical imaging [31, 32], virtual reality [33], face detection [34, 35], cultural heritage [36], and quality control [37]. The aim is to recognize $3D$ objects by relying on their shapes' properties extracted by the mean of a shape descriptor. In fact, the recognition process is performed in three steps. The first step consists of computing descriptors, which can be global [1], local [2], structural [11, 12], transform-based [55, 59], or view-based [38, 56]. In the second step, the similarity measure between the $3D$ shapes through their associated descriptors is performed. It can be distance-based [17], probability-based [18], or graph-based [48]. Relying on the results of the similarity measure performed in the second step, the $3D$ objects are classified via a decision process, that can be hard [29, 30], or soft [30, 50].

In this work, the topological invariants provided by the critical points of the $3D$ object and the link between them are exploited. The main idea of this description is to represent one or several features in a real function, to define it on the $3D$ object, and to describe its topological changes by the critical points and the links between them. In fact, some algebraic topology tools [6, 7, 16] express these topological changes by counting the number of connected components, the number of tunnels, and the numbers of holes between critical points in the $3D$ object. In addition, these tools require the fact that the features be represented in a function that satisfies Morse constraint [16]. Besides, if a $3D$ object underwent a cut or erosion process, undesired holes would be added to the number of holes to be computed, with the number of connected components and the number of tunnels. To overcome these shortcomings, the description process will be only based on counting the number of connected components between the critical points in order to describe incomplete $3D$ objects. For that purpose, we chose to describe the $3D$ object by the tool of size function. Its principle is to encode quantitatively the topological changes provided by the critical points of the $3D$ object [16]. In fact, it is resulted from a real continuous function defined on the $3D$ object called measure function [11], with the aim to capture some specific features of the $3D$ object [10]. Unlike the related methods in the literature [6, 7], the advantage of describing $3D$ objects by size functions is that the measure function can or cannot be

a Morse function [16]. The topological changes have been used for 2D and 3D object recognition [8, 9, 10, 12].

This paper is organized as follows: in Section 2, some 3D object recognition methods of the related work are presented and discussed. Our proposed method is described in Section 3. The interpretation of the obtained results using the proposed method and some methods in related work is presented in Section 4. Finally, conclusions and future works are presented in Section 5.

2.2 Related work

The recognition of a 3D object can be performed by the comparison of a 3D object to several 3D objects, several 3D objects to a 3D, or several 3D objects to several 3D objects [50]. After the comparison, the 3D objects are classified with the help of a decision process, which can be hard [29, 30] or soft [30, 50]. In fact, when the decision is hard, the 3D object can resemble one of the 3D objects. The principle of the soft decision is to sort the 3D objects according to their resemblance to the 3D object to be recognized. In order to identify similar 3D objects, several strategies have been proposed to perform this task such as the nearest the ranking [50], neighbours [50], the Bayesian decision rule [60], the risk functions [49], and the discrimination [52]. The choice of the criterion of similarity and decision depends on the 3D object collection's size and the feature space [50, 51].

The 3D object recognition methods can be categorized into five approaches [20, 21] which are local, global, transform-based, views-based, and structural. In the following paragraphs, the different groups of approaches, the description principle of each one of them, the similarity measure, and the used decision process are described.

The aim of the local approaches is to describe the 3D object based on the neighbourhood of each vertex on it. In the work of Mousa [2], the 3D object is described by its principle curvatures. For each vertex of the 3D object, principle curvatures are computed and stored in a matrix. The first column of the matrix contains the minimum curvatures, and the second column contains the maximum curvatures of the 3D objects. This matrix is the descriptor of the 3D object. The similarity measure between the 3D objects is performed by using the Euclidian distance between the

matrices associated with the $3D$ objects, and the decision is soft. Xiaofeng et al. [5] described the $3D$ object by the projected areas of its mesh's vertices. At first, the projected area on the vertical plane of the normal vector is computed for each vertex of the mesh that represents the $3D$ object. Afterwards, the feature vector is formed is computed by transferring the projected areas to the $1D$ Fourier transform to obtain the integral invariants. Finally, the Euclidean distance between feature vectors associated with the $3D$ objects is used to compare between $3D$ objects, and the decision is soft.

The principle of the global approaches is to describe the $3D$ object based on its geometric properties. It can be carried out by computing: for example, a histogram associated with a specific feature of the $3D$ object, or distances between vertices of its associated mesh. In the work of Ion et al. [1], the $3D$ object is described by an eccentricity histogram of its vertices. The Euclidian distance between the associated eccentricity histograms of the $3D$ objects is used as a similarity measure, and the decision is soft. A global descriptor that contains geodesic information of the mesh representing the $3D$ object is used by Rabin et al. [3] to characterize the $3D$ object. Before computing the global descriptor, the triangular mesh that represents the $3D$ object is sampled into 500 vertices. Afterwards, the geodesic distance, the median geodesic distance, and the eccentricity transforms are computed for each vertex of the $3D$ object to form the global descriptor associated with the $3D$ object. Finally, the distance of Wasserstein [61] between the global descriptors associated with the $3D$ objects is used as a similarity measure, and the decision is soft.

In the transform-based approaches, the descriptor associated with the $3D$ object is formed by some coefficients of high frequency that contain most of the energy in the transform domain, especially when the transform is orthogonal. In the work of Laga et al. [59], the spherical wavelet transform is used to extract three $3D$ object descriptors. At first, the $3D$ object is normalized using the principal component analysis [15] to ensure the features invariance to affine transformations. Afterward, the first feature vector is extracted from a subset of spherical wavelet coefficients. The spherical wavelet sub-bands' energies L_1 and L_2 are computed to form the two other feature vectors. Finally, the $L_2 - distance$ between the associated feature vectors of the $3D$ objects is used as a similarity measure, and the decision is soft.

The principle of the view-based approaches is to describe the 3D object by a set of 2D descriptors associated with its set of 2D views. In the work of Su et al. [56], the 3D object is described by a set of 2D rendered views. In fact, the authors presented a multi-view convolutional neural network architecture (CNN). The principle is to train the CNN associated with each view independently and to combine information from all the views of the 3D object into a single shape descriptor. In order to endow the descriptor invariance to rotation and to translation, transformed copies of the 3D objects are added during the training process. A well-suited distance is used to perform the similarity measure between the shape descriptors associated with the 3D objects, and the decision is soft.

The main idea of the structural approaches is to describe the 3D object by relying on the analysis of its topological changes based on a function defined on it. Baloch et al. [11] described the 3D object by a topo-geometric shape model. The principle is to represent the 3D object by a skeletal graph that is resulted from a Morse function. The similarity measure between 3D objects is performed by measuring the subgraph isomorphism between their associated skeletal graphs, and the decision is soft. In the work of Biasotti et al.[12], the 3D object is described by a multidimensional size function. The principle is to represent the 3D object by a size function resulting from a multidimensional measure function. The aim is to encompass several measure functions, where each represents some specific features of the 3D object in a multidimensional measure function defined on the 3D object. To reuse the 1D size functions' principles, a theorem is conceived that generalizes the 1D size functions' concepts for the multidimensional case. The authors developed a lower bound matching distance to compare between the multidimensional size functions associated with the 3D objects, and the decision is soft. In our early work [28], the 3D object is described locally by a set of 18 size functions. At first, the 3D object is normalized by ISOMAP [14] and PCA [15] in order to endow the features to be extracted invariance to articulations, torsions and affine transformations. Afterwards, the 3D object is split into a set of 18 portions, where each one is represented by a size function resulted from the concavity and convexity features. A ramified average minimum distance between the set of 18 size functions is used to compare between the 3D object, and the decision is soft.

Table 2.1 represents a summary of the mentioned methods for describing 3D

Table 2.1 – Summary of methods for the recognition of 3D objects.

Approaches	References	Descriptors	Types of representation	Invariances	Similarity measures	Decision	3D Objects collection	Computer configuration
Global	[1]	Eccentricity histogram	Mesh and volume	Affine transforms, articulations, andtorsions	Euclidian distance	Soft	255 3D objects of the McGill Shape Benchmark	-
	[3]	Global descriptor of geodesic information	Mesh	Affine transforms, articulations, andtorsions	Distance of Wasserstein	Soft	457 3D objects of the McGill Shape Benchmark	-
Local	[2]	Principal curvatures	Mesh	Affine transforms, articulations, andtorsions	Euclidian distance	Soft	3D objects from the Web	PC with a dual-core 2.8GHz processor and 2GB of memory
	[71]	Projected areas of vertices	Mesh	Affine transforms	Euclidian distance	Soft	867 engineering models of ESB 3D object collection	
Transform-based	[6]	3D Fourier descriptor	Mesh	No	Euclidian distance	Soft	1830 3D objects from 3Dcafe and viewpoint sites	850 MHz Processor Pentium III, windows 2000
	[59]	Spherical wavelet coefficients, L_1, and L_2 energies of the spherical wavelet sub-bands	Mesh	No	Euclidian distance	Soft	1814 3D objects of the Princeton Shape Benchmark	-
View-based	[38]	49D Zernike moments associated to the 41 2D views	Mesh	Affine transforms	Kullback-Leibler divergence	Soft	The 3D object collections of ETH, NTU and Princeton Benchmark	-
	[56]	Multi-view convolutional neural network descriptor	Mesh	Translations and rotations	Distance between shape descriptors	Soft	12311 3D objects of the Princeton ModelNet dataset	-
Structural	[11]	Topo-geometric shape model	Mesh	Affine transforms	The matching is based on measuring the subgraph isomorphisms between skeletal graphs	Soft	3D objects collection of GAMMA provided by INRIA	-
	[12]	Multi-dimensional size functions	Mesh, and volume	Affine transforms, articulations, andtorsions	Lower bounds matching distances	Soft	255 3D objects of the McGill Shape Benchmark	AMD Athlon 3500, 2 GB RAM
	[28]	Set of 18 size functions	Mesh	No	Ramified average minimum distance	Soft	2902 3D models of McGill, architectural data [64], 3D objects from Robert Sumner's collection of 3D shapes [44], and 3D models from the Princeton's collection of 3D objects [62].	PC with 2.80 Ghz Intel Xeon® CPU with RAM 12.2 Go and NVIDIA Tesla c2050/c2070 GPU computing processors, and running under Windows 7 64 bits

objects, which belong to the five approaches of 3D object recognition. According to Table 2.1, most of the description methods rely on the triangular mesh that represents the 3D object. The choice of descriptors presented in Table 2.1 is related to the collection of the 3D objects, which has categories that may contain some specific problems. For example, the structural descriptors defined in [11, 12] and the global descriptors [1, 3] showed invariance to affine transformations, to articulations, and to torsions when they were validated using the articulated 3D models of the 3D object collection of McGill [19]. We noticed that only 255 3D objects of the 3D objects in the collection of McGill were used to validate some descriptors [1, 11, 12]. Besides, the description process for all the mentioned methods did not include contextual information of the 3D objects, because the used collections of 3D objects contain isolated 3D models. The Euclidian distance was used as a similarity measure between 3D objects by most description methods. However, the similarity measure used by the structural descriptors is graph-based because the features are represented in graphs.

Regarding the decision, the retrieval process is performed by using a soft decision for all the mentioned descriptors.

The shortcomings of our early work [28], are that the measure function is not invariant to articulations, torsions and affine transformations. In addition, the splitting criterion is sensitive, because it supposes that the $3D$ objects are complete, which is not always the case. Thus, the local approach proposed in [28] is not well suited to the problem of the partial matching of the $3D$ objects. In addition, the ramified distance used as a similarity measure considers that the matchings have the same contribution. Therefore, the use of irrelevant matchings in the comparison process of $3D$ objects can add confusion between $3D$ objects of different categories during the classification process. For that purpose, the principle of our method is to describe the $3D$ object by a size function resulting from a measure function that captures the surface displacement feature along the $3D$ object, and that encompasses invariance to affine transformations, articulations, and torsions. To overcome the problem of occlusion and noise, a well-suited metric learning method is used to adjust the matchings according to their relevancies. In the next section we will further the principles of size functions in the context of $3D$ object recognition.

2.3 Proposed method

The size function has been developed since the beginning of the 1990s by Frosini et al. [58] and Verri et al. [39]. It was used for the recognition of sign language [9], automatic classification of melanocytic lesions [45], automatic classification of white blood [4], and $3D$ object recognition [12, 28]. The principle of size function is to describe the $3D$ object by encoding the topological changes provided by its critical points and the links between them. Indeed, the critical points are those of a real function defined on the $3D$ object, called measure function; which has values on its critical points called critical values. A critical point of the measure function defined on the $3D$ object can be a local or a global maximum or minimum or saddle point. The main idea is to represent a feature in a function to define it on the $3D$ object, and to compute the critical points and the links between them to describe the $3D$ object. The measure function can be, for example, a distance between the vertices of $3D$ object,

or a function that represents curvature on each vertex of it, etc. In size function, the critical points have physical significations, such as the appearance or disappearance of one or several connected components. In fact, the critical points are computed by a topological analysis of the measure function, as will be explained in Section 3.1. To illustrate, we consider the cyclide in Figure 2.1.(a) as defined by a function φ that takes values in the R axis. The critical points P_1, P_2, P_3, and P_4 associated with the critical values a, b, c, and d, indicate the appearance or the disappearance of one or several connected components. For example, between the critical values a and b associated with the critical points P_1 and P_2, there is only one connected component, between the critical values b and c associated with the critical points P_2 and P_3, there are two connected components, and so on.

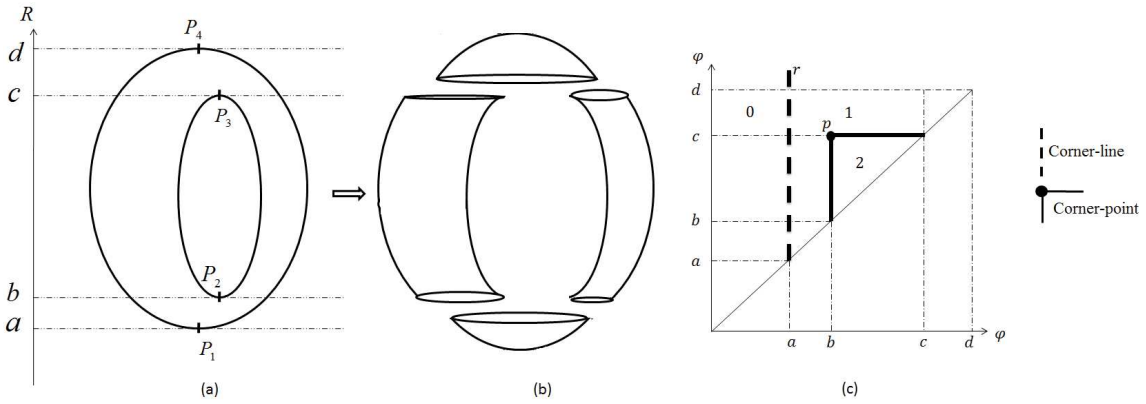

Figure 2.1 – (a) The critical points, (b) the connected components, (c) and the resulting size function.

Besides, if a measure function ψ that takes values in the R' axis is defined on the same cyclide, it will be represented by other critical points, and consequently by other connected components as illustrated in Figure 2.2. Thus, the definition of different measure functions in the same $3D$ object implies different descriptions.

2.3.1 $3D$ object description

Let M be a connected topological space endowed by a continuous function φ defined as follows: $\varphi : M \to \mathbb{R}$. The function φ is the measure function, and the pair

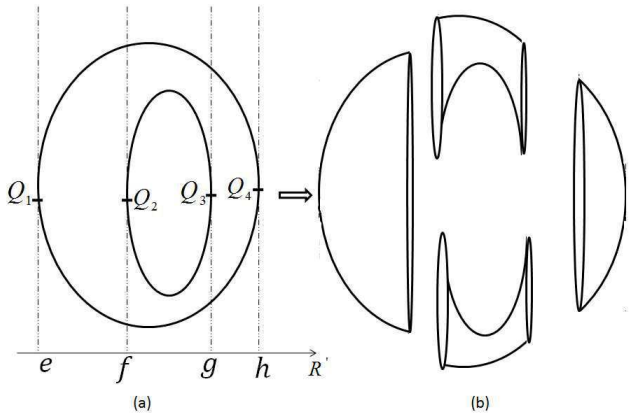

Figure 2.2 – (a) The critical points, (b) and the connected components according to the R' axis.

(φ, M) is called a size pair. The size function $l_{(M,\varphi)}(x, y)$ corresponds to the number of connected components of the subset $\{P \in M, x < \varphi(P) < y\}$ and it is defined as follows: $l_{(M,\varphi)} : \mathbb{R} \times \mathbb{R} \to \mathbb{N} \cup \{\infty\}$. Let consider the cyclide example illustrated in Figure 2.1.(a). In fact, when $a < \varphi(P) < b$, there is only one connected component, the size function takes the value one. When $b < \varphi(P) < c$, there are two connected components illustrated on Figure 2.1.(b), the size function takes the value two, and so on. The outcome is the function graph represented in Figure 2.1.(c), in which the regions where the size function takes different constants are demarcated by corner-points. In Figure 2.1.(c), the bold dashed line r is a corner-line, and it denotes the appearance of the first critical point associated with the global minimal critical value taken by the measure function φ. The point p is a corner-point that represents the intersection point between the two bold solid lines, which demarcate the region where the size function takes the value one, and the region where the size function φ takes the value two. From Figure 2.1.(b), the corner-point $p = (b, c)$ represents a detail of the $3D$ object, and it is associated with the local minimum critical point P_2 where the measure function takes the critical value b, where two new connected components appear, and the saddle critical point P_3 where the measure function takes the critical value c, where the two connected components disappear.

The aim is to represent the $3D$ object by a measure function on its mesh. In fact, if the positions, poses, articulations and torsions of a $3D$ object change, then

the inter-vertices Euclidean distance changes. However, the surface of the $3D$ object does not change. Therefore, the displacement from a vertex P_i to another vertex P_j along the $3D$ object's mesh does not change, and it can be expressed by the geodesic distance between the two vertices P_i and P_j. For that purpose, a measure function that relies on the surface displacement along the mesh that represents the $3D$ objects will be invariant to translations, rotations, articulations and torsions, and consequently, it will preserve the topological invariants provided by its critical points on the $3D$ objects. For that aim, we define the measure function φ for each vertex P_i of the mesh M as follows:

$$\varphi(P_i) = \frac{f(P_i)}{\max\limits_{P_j \in M} f(P_j)} \tag{2.1}$$

Where the function $f(P_i)$ is defined as follows:

$$f(P_i) = \sum_{j=1}^{N} e^{-\frac{d_{geodesic}(P_i,P_j)}{\max\limits_{P_k,P_l \in M} d_{geodesic}(P_k,P_l)}} area(P_j) \tag{2.2}$$

Where N is the number of vertices in the mesh M, $d_{geodesic}(P_i, P_j)$ is the geodesic distance between the vertices P_i and P_j and it is computed by finding the shortest path along the mesh's graph that connects the two vertices P_i and P_j with the help of Dijkstra's algorithm [25], and $area(P_j)$ is the sum of triangles' surfaces that touch the vertex P_j. The role of the denominator inside the exponent in the function f is to ensure the term of geodesic distance invariance to scale. The denominator in the measure function φ endows its invariance to scale, because the surface of the $3D$ object changes with the scale's variation. In fact, the measure function φ determines how displacement from a vertex P_i to a vertex P_j is performed along the surface of the $3D$ object, according to its topology. The examples of two articulations of the letter $'n'$ in Figure 2.3 illustrate the critical points of the measure function φ are preserved for the letter $'n'$ even if with the change of articulations. Thus, the same connected components are obtained as shown in Figure 2.3.(c), and consequently, the resulting size functions are the same as illustrated in Figure 2.3.(d).

Besides, the measure function's graph often has a large number of vertices, which

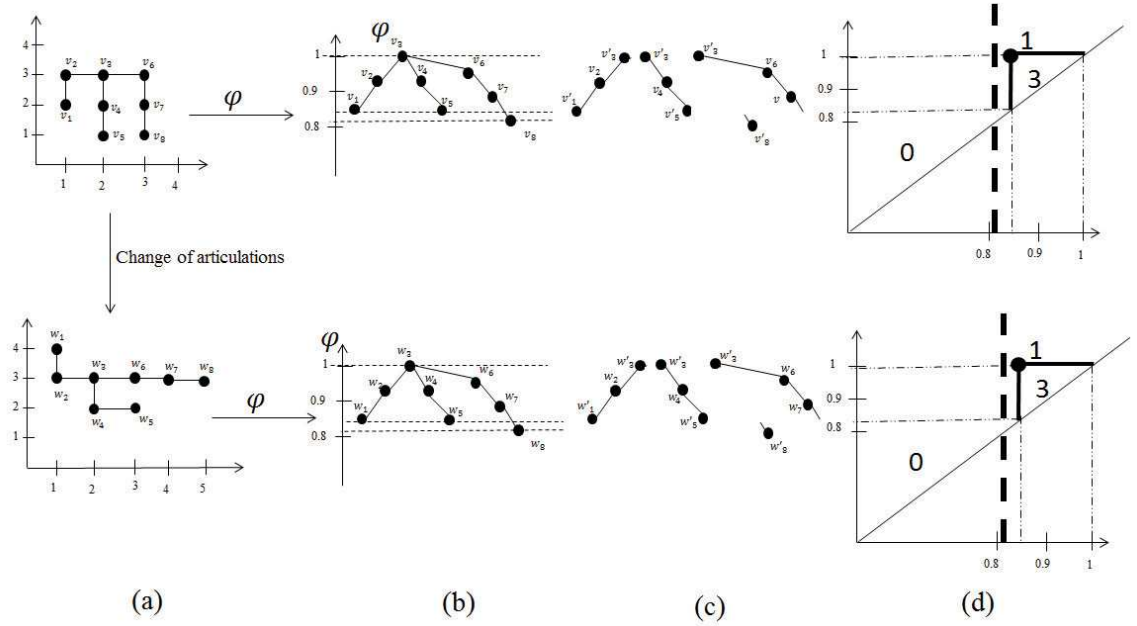

Figure 2.3 – (a) Two articulations of the letter $'n'$, (b) representations of the measure functions, (c) connected components, (d) and the resulting size functions.

implies a large storage capacity in memory, and makes the computation of size function more time-consuming. To overcome these shortcomings, the $3D$ object can be sampled in a small set of vertices, before computing the measure function. However, some details of the $3D$ object that can be relevant for its description may be lost because of this sampling process. In order to avoid this problem, a simplification process of the measure function's graph is performed with the help of a $\Delta^* - reduction$ algorithm [23, 27]. The principle is to reduce the measure function's graph to obtain an oriented arborescence that contains only the critical points and the links between them. Thus, the corner-lines and the corner-points of the resulting size function are computed directly from this arborescence with the help of D'Amico's algorithm [23].

2.3.2 Similarity measure

As mentioned previously, each $3D$ object is described by a size function. Thus, the comparison between two $3D$ objects is performed through the similarity mea-

sure between their associated size functions. Besides, the size function is represented by a graph composed of a finite number of links between critical points, which are corner-lines, and corner-points in the upper diagonal. For that purpose, the similarity measure between two size functions is graph-based, and it can be performed by computing the correspondence between their associated graphs. D'Amico et al. [13] developed a matching distance, which principle is to measure the maximum cost of displacing the corner-lines and corner-points of a size function to those of the other one as illustrated in Figure 2.4. However, if the $3D$ object is affected by noise or occlusion, some or all of its critical points and the links between them change, and consequently its description changes. Therefore, the matching distance proposed in [13] has a limitation, since the matchings based on the comparison between corner points and corner lines that are generated from noise and occlusion have the same contribution as the relevant ones. To overcome this shortcoming, a metric learning process must be performed in order to take into account the contribution of each matching by weighting it. The principle is to increase the effects of matchings that facilitate the identification of $3D$ objects that belong to the same category to reduce the mismatching. For that purpose, we distinguish two cases that imply two different processes. The first case is when the $3D$ objects belong to the same category. The effects of matchings between the corner-points and corner-lines that are generated from noise and occlusion must be decreased. In the second case, the $3D$ objects belong to different categories. Indeed, the contributions of matchings that add confusions between $3D$ objects must be decreased, and the effects of matchings that facilitate the discrimination between $3D$ objects must be increased. Consequently, the problem is the selection of the useful matchings between the $3D$ objects. To do so, a weighting process must be performed. Let N be the number of $3D$ objects in the collection of $3D$ objects, and $S = \{S_1, ..., S_N\}$ the set of similarity vectors between the $3D$ objects of the collection of the $3D$ objects. In fact, these similarity vectors are composed of the displacements between corner-points and corner-lines. To generate each similarity vector S_i, let consider two $3D$ objects O_1 and O_2 described respectively by size functions l_1 and l_2, where their associated corner-lines and corner-points are stored respectively in the sets C_1 and C_2 in the ascending order of their abscissas. Each element of S_i is generated by computing the matching distance between the closest corner-

points/lines according to their abscissas as shown in the example in Figure 2.4. In general, if the size function l_2 has an inferior number of corner-lines and corner-points than those of the size function l_1, its set of corner-points is completed by the orthogonal projection of the remaining corner-points of the size function l_1 at the diagonal, as illustrated in the example of Figure 2.4.(c). Meanwhile, the similarity vectors may have different dimensions because $3D$ objects have different corner points and corner lines. To overcome this shortcoming, each similarity vector is completed by zeros to have the same dimension as the similarity vector that has the maximum dimension. Indeed, the matching distance's principle of the equation (3) relies on computing the minimum displacement of the corner point to the other one, and the displacement of the two corner points to the diagonal.

$$\delta\left(p,p'\right) \qquad = \qquad \min\left\{\max\{|x-x'|,|y-y'|\},\max\left\{\frac{|y-x|}{2},\frac{|y'-x'|}{2}\right\}\right\} \quad (2.3)$$

The obtained similarity vectors are classified into two groups. The first group is the intra-classes similarity vectors, which are obtained by measuring the matchings between $3D$ objects of the same category. In the second group, which is called inter-classes similarity vectors, the matchings between $3D$ objects of different categories is measured. The aim is to separate the intra-classes group of similarity vectors and the inter-classes. For that purpose, a weighting process of the matchings stored in the similarity vectors is performed using a multi-layer neural network [24]. For our case, we used a four-layer neural network, where the inputs are the matchings stored in the similarity vector. For the two hidden layers, we set dimensions of $2N$ for the first hidden layer, and $N/2$ for the second hidden layer where N is the dimension of the similarity vector. In fact, we used a neural network with two outputs, where the first one contains the rate that the similarity vector belongs to intra-classes group, and the second output contains the rate that the similarity vector belongs to inter-classes group. Before proceeding to the learning process, the desired output is set. Indeed, if the input vector is obtained from the matching between $3D$ objects that belong to the same category, the first output value is set to one and the second value is set to -1. Otherwise, the first output value is set to -1 and the second value is set to one. The learning process is performed with the help of the algorithm of short for resilient

backpropagation [22].

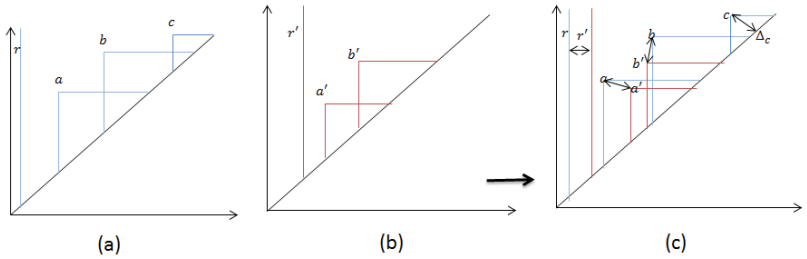

Figure 2.4 – Example of matching between two size functions (a) l_1 with $C_1 = \{r, a, b, c\}$ and the size function (b) l_2 with $C_2 = \{r', a', b'\}$, given by $d_{match}(l_1, l_2) = \max\{\delta(r, r'), \delta(a, a'), \delta(b, b'), \delta(c, \Delta_c)\}$.

2.4 Experimental results

At first, the proposed method was experimentally evaluated using the manifold version of the $3D$ objects collection of McGill [19]. The latter is composed of 457 $3D$ objects in high-level semantics 19 categories, where the $3D$ objects of the same category have different shapes, articulations, torsions, positions, scales, and orientations. The 255 $3D$ objects in the first ten categories of this $3D$ objects collection have articulated parts; the remaining nine categories contain $3D$ objects with non-significant articulated parts. Figure 2.5 illustrates examples of $3D$ objects of each category of the 19 classes of the $3D$ objects collection of McGill. In fact, it is one of the standard $3D$ models collections used to validate several descriptors of $3D$ objects [1, 3, 11, 12]. In order to assess our method's performance in terms of partial matching, we evaluated it by using the $3D$ objects collection proposed by Rodola et al. [54]. In fact, the selected directory of this $3D$ object collection contains two sets of $3D$ objects. The first set is called cuts, where the $3D$ objects were passed through a single cut. In the second set, which is called holes, the $3D$ objects contain multiple holes and were subjected to multiple cuts. The holes were generated by applying an erosion process to the surface of the $3D$ object. The effect of holes on each $3D$ object is expressed by a value that indicates the fraction of kept area of the $3D$ object after the erosion process. In addition, there is a directory called null that contains 76 $3D$ objects with

no cuts nor holes, but with different articulations and torsions. The total number of 3D objects collection from the selected and null directories is 275 and they are attributed to eight categories. Figure 2.6 illustrates examples of 3D objects of the eight categories of the 3D objects collection of Rodola et al. [54]. Afterwards, the proposed method was experimentally evaluated by merging different 3D objects collections. For that purpose, we repaired 3D objects from the 3D shape benchmark of architectural data [64] and the Princeton 3D objects collection [62] with the help of the algorithm proposed in [63]. The obtained 3D objects were added with the articulated 3D objects of Robert Sumner's shapes repository [44] to the 3D objects collection of McGill. The outcome is a 3D object collection of 2902 objects and 121 categories.

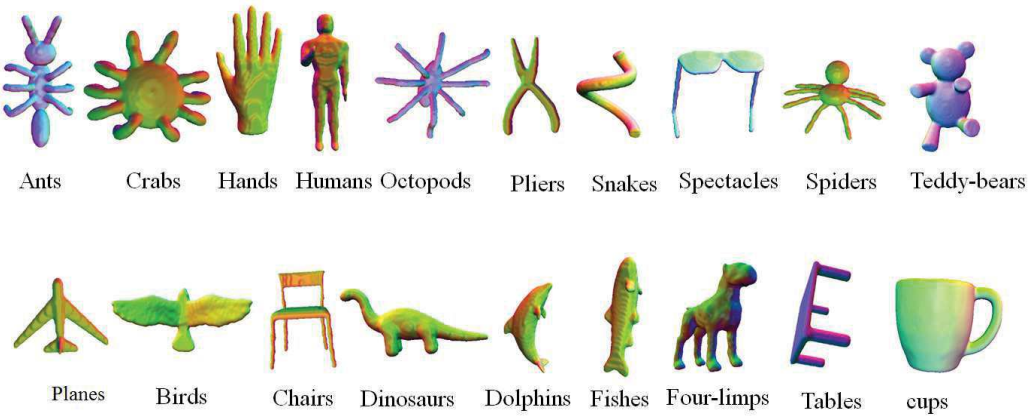

Figure 2.5 – Examples of 3D objects from the 3D objects collection of McGill (Top) with articulated parts (Bottom) and non-significant articulated parts.

Our 3D objects recognition system is summarized in Figure 2.7. At first, each 3D object is described by a size function that will be stored in the features dataset. Afterwards, a metric learning process using neural networks is performed. Thus, the comparison between the 3D objects is obtained through the prediction relying on metric learning process parameters. Finally, a soft decision is performed in order to retrieve the 3D objects according to their resemblance to the 3D query object.

To evaluate the proposed method, each 3D object is used as a query object, and a soft decision process is performed to sort the 3D objects of the collection according to their resemblance to the 3D object to be recognized. The recognition performance was evaluated in terms of recall and precision [26], nearest neighbour, first tier, and

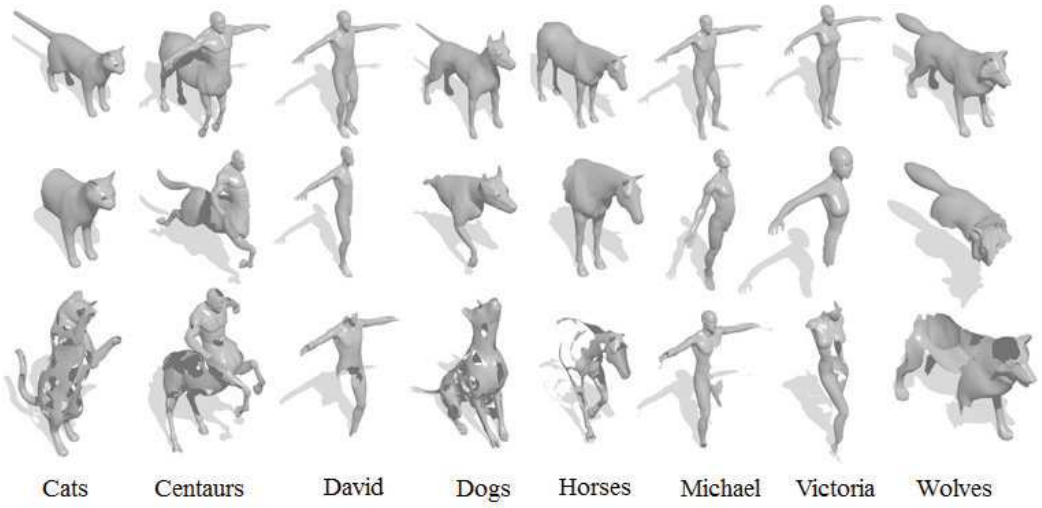

Figure 2.6 – Examples of 3D objects from the 3D objects collection of Rodola et al. [54] from (Top) null, (middle) cuts, (Bottom) and holes directories.

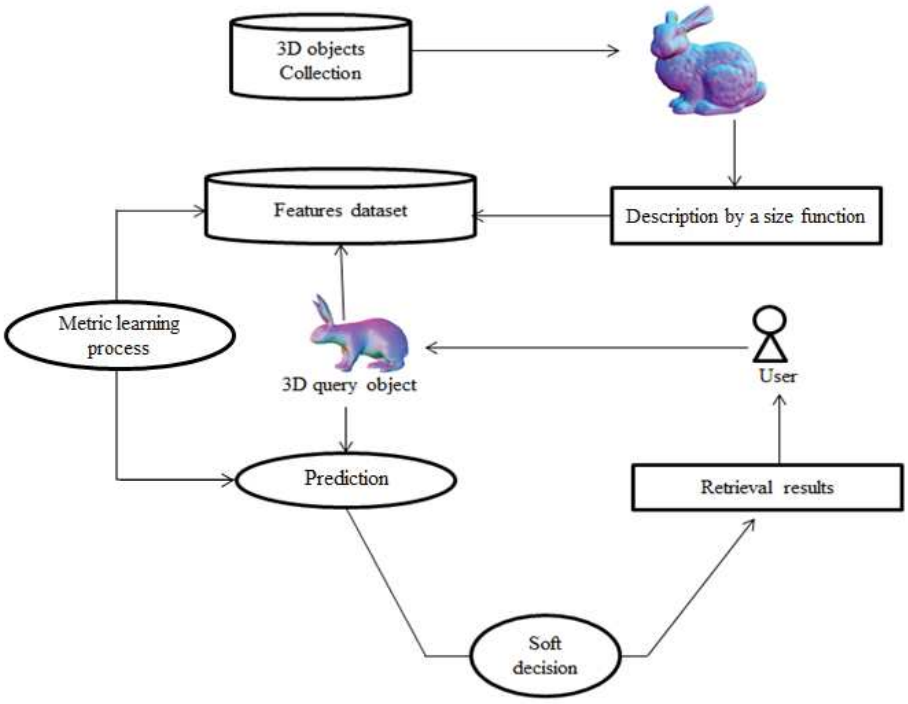

Figure 2.7 – The 3D object recognition system.

the second tier evaluation measures [19]. For that purpose, a query is formed by a $3D$ object belonging to a category C_i of Q $3D$ objects. The recall represents the number of $3D$ objects belonging to the category C_i from the k first retrieved $3D$ objects, we denote it by $N_{corrects}(k)$, on the number Q of the $3D$ objects of the category C_i. The precision is the number of $3D$ objects that belong to the category C_i found from the k first retrieved $3D$ objects, on the number of responses k. The precision and recall can be formulated as follows:

$$Precision(k) = \frac{N_{corrects}(k)}{k}, \tag{2.4}$$

$$Recall(k) = \frac{N_{corrects}(k)}{Q} \tag{2.5}$$

The nearest neighbour (NN) measure provides the recognition score of the first retrieved 3D objects that belong to the same category of the 3D object used as a query. The first tier has the same principle of the recall because it provides the recognition score, where the number of retrieved $3D$ objects is the number of $3D$ objects in the category of the $3D$ query object. The second tier provides the recognition score, where the number of retrieved $3D$ objects is twice the number of $3D$ objects in the category of the $3D$ query object.

Table 2.2 – The average recognition scores of different methods for the 255 articulated $3D$ objects collection of McGill.

Methods	Nearest-neighbor	First tier	Second tier
Our method with metric learning	98.82%	**81.57%**	**91.47%**
Our method using distance in [13]	**100.00%**	67.04%	78.22%
Tabia et al. 2015	97.70%	73.20%	81.80%
Local approach. 2014	**100.00%**	43.69%	62.45%
Tabia et al. 2013	96.90%	65.80%	78.10%
Lavoué 2012	95.70%	63.50%	79.010%
Agathos et al. 2009	97.60%	74.10%	91.10%
Papadakis et al. 2008	92.50%	55.70%	69.80%

Table 2.2 summarizes the scores obtained using the first 255 $3D$ objects in the first ten categories of this $3D$ objects collection of McGill by our method with metric learning, our method using distance in [13] the work of Tabia et al.[40], where the $3D$ object is described by a covariance descriptor, a local approach [77] where the $3D$ object is split into 18 pieces to describe each one of them by a size function, the work of Tabia et al.[42] where the $3D$ object is described by vectors of locally aggregated tensors, the work of Lavoué [43] where the principle of Bag of Words is used to describe the $3D$ object, the work of Agathos et al.[41] where the attributed relational graph is used to describe the $3D$ object, and the work of Papadakis et al. [53] where a compact hybrid shape descriptor is used to represent the $3D$ object. We took the scores from the paper of Tabia et al. [40]. For the learning process, we used the 255 $3D$ objects in the collection of McGill. As illustrated on Table 2.2, our method with metric learning is favourably comparable in terms of nearest neighbour with the related work, except for the local approach [28] and our method using distance to [13] where the nearest neighbour reaches 100%. Besides, in terms of first and second tiers our method with metric learning outperforms all the methods of the related work. In fact, the metric learning improved our method's recognition scores to overcome those obtained on related work.

We compared the proposed method with the work of Biasotti et al.[12] where they described a $3D$ object by a multidimensional size function, the work of Ion et al.[1] where the $3D$ object is described by an eccentricity histogram, and the local approach of Xiaofeng et al.[5] where the $3D$ object is described by the local areas of its vertices. For the learning process, we used the 457 $3D$ objects in the collection of McGill, 275 $3D$ objects collection proposed by Rodola et al. [54], and 1279 $3D$ objects of the extended collection as a learning data. We implemented the mentioned methods in C++ programming language.

Figure 2.8 and Figure 2.9 represent respectively the average precisions/recalls curves of our method with metric learning, our method using the distance in [13], local approach [28], the work of Biasotti et al. [12], the work of Ion et al. [28], and the local approach of Xiaofeng et al. [5] for the 457 $3D$ objects of the McGill $3D$ models collection, and for the 2902 $3D$ objects of the extended collection. Table 2.3 summarizes the accuracy of our method, the scores obtained by the method Zimmer

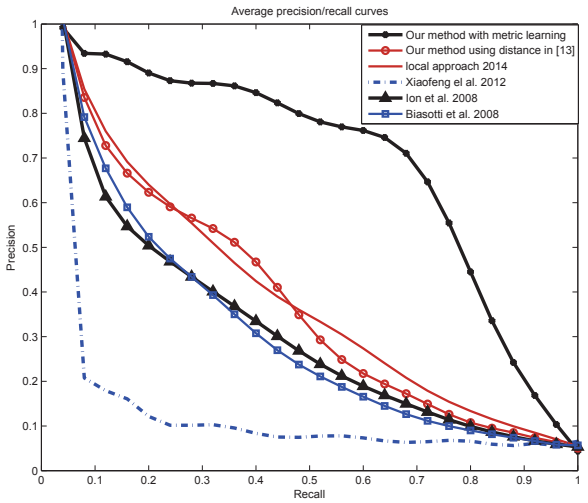

Figure 2.8 – The averages precisions/recalls curves for the 457 3D objects collection of McGill.

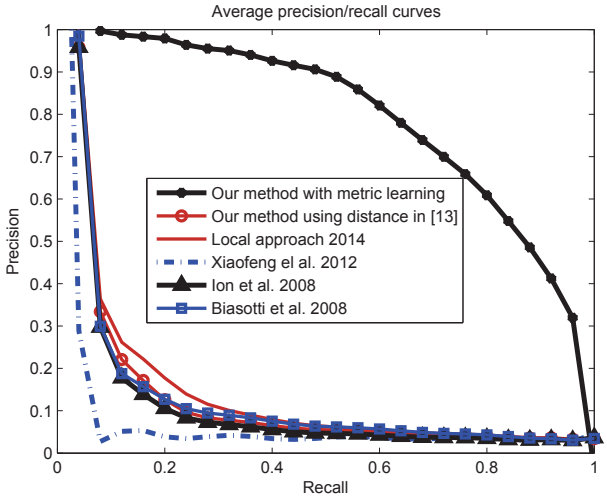

Figure 2.9 – The averages precisions/recalls curves for the 2902 3D objects of the extended collection.

et al. [57], the scores obtained by the method of Lian et al. [47], and the other approaches for the 457 $3D$ objects collection of McGill by the means of the first and second tier evaluation measures [19]. Table 2.4 summarizes the accuracy of our method, and the other approaches for the 2902 $3D$ objects of the extended collection by the means of the first tier, and the second tier. As illustrated on Figure 2.9 and Table 2.3, our method got the highest scores when it was compared with the other approaches of the related works for the 457 $3D$ objects collection of McGill. From Figure 2.9 and Table 2.4, the scores obtained by our method with metric learning are higher than those of the related works for the 2902 $3D$ objects of the extended collection. On the one hand, even if the $3D$ objects of the same category have different articulations, torsions, positions, scales and poses, the measure function that relies on the surface displacement between mesh's vertices preserved the topological invariant provided by its critical point on the $3D$ object. On the other hand, the weighting of the matchings between $3D$ objects obtained by the metric learning enhanced our method's performance to get more than 93% as a recognition rate.

Table 2.3 – The average recognition scores of different methods for the 457 $3D$ models of $3D$ objects collection of McGill.

Methods	Nearest-neighbor	First tier	Second tier
Our method with metric learning	98.47%	**66.52%**	**79.01%**
Our method using distance in [13]	**100%**	41.76%	55.43%
Local approach 2014	**100%**	40.88%	57.17%
Zimmer et al. 2013	51.20%	26.6%	40.0%
Lian et al. 2012	45.9%	24.9%	38.5%
Xiaofeng et al. 2012	100%	14.98%	22.81%
Ion et al. 2008	100%	35.70%	50.37%
Biasotti et al. 2008	100%	35.42%	48.24%

Figure 2.10 and Table 2.5 represent, respectively, the average precisions/recalls curves and the scores obtained by the means of the nearest neighbour, first and second tiers of our method with metric learning, our method using the distance in [13], and three previously mentioned methods for the 275 $3D$ objects collection proposed

Table 2.4 – The average recognition scores of different methods for the 2902 3D objects of the extended collection.

Methods	Nearest-neighbor	First tier	Second tier
Our method with metric learning	**100%**	**93.21%**	**96.15%**
Our method using distance in [13]	99.93%	13.30%	16.78%
Local approach 2014	99.93%	13.24%	18.18%
Ion et al. 2008	**100%**	11.11%	14.59%
Biasotti et al. 2008	99.93%	10.77%	14.35%
Xiaofeng et al. 2012	91.80%	5.36%	6.83%

Table 2.5 – The average recognition scores of different methods for the 275 3D objects collection proposed by Rodola et al. [54].

Methods	Nearest-neighbor	First tier	Second tier
Our method with metric learning	**100%**	**79.56%**	**86.79%**
Our method using distance in [13]	**100%**	17.77%	29.60%
Local approach 2014	**100%**	24.02%	37.83%
Ion et al. 2008	**100%**	29.87%	46, 86%
Biasotti et al. 2008	**100%**	18.09%	31, 26%

by Rodola et al. [54] for partial matching. In fact, despite the fact that the 3D objects of Rodola et al. [54] were subjected to multiple cuts and contain several holes; our method with metric learning got the highest scores. Figure 2.11 and Table 2.6 illustrate that the recognition rates of our method are always higher than those of the related work; and they remained higher than 77%, even if some 3D objects of the collection of Rodola et al. [54] lost 60% of their areas. Thus, the metric learning permitted the use of the most relevant matchings to identify 3D objects of the same category, and to distinguish between 3D objects of different categories. Figure 2.12 illustrates some examples for retrieving 3D objects by using a query of complete 3D object and only a portion of 3D object. However, our method using the distance

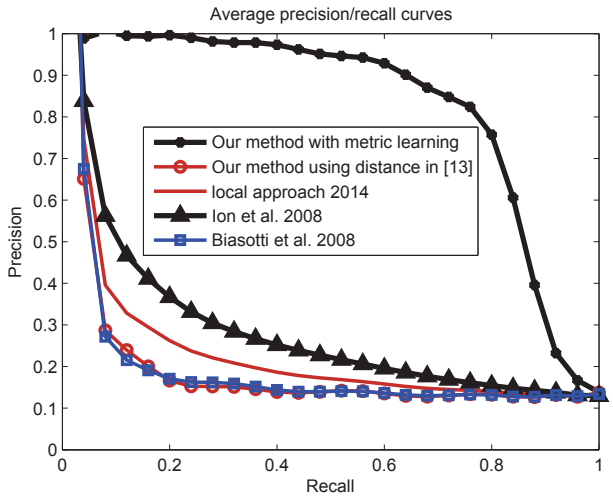

Figure 2.10 – The averages precisions/recalls curves for the 275 $3D$ objects collection proposed by Rodola et al. [54].

in [13], the local approach of size function [28], and the work of Biasotti et al. [12] got the lowest scores. On the one hand, the principle of the local approach of size function described in [28] is to split the $3D$ object into 18 portions to represent each one by a size function. However, the splitting process is based on the three principal axes of the $3D$ objects. Thus, if the $3D$ object has missing parts or holes, its three principal axes change, and consequently, its description changes. On the other hand, the matching distance described in [13] and the matching distance used by Biasotti et al. [12], have the same principle to compute the maximum displacement between the corner points/lines of size functions. Besides, when the $3D$ object has holes, or underwent cuts, the critical points and the links between them change, and therefore, its description changes. As a result, the matching distances relying on the computation of the maximum matching may select a matching that is not useful to compare between $3D$ objects.

Figure 2.13 represents the average computation time of the different methods obtained on a PC with 2.80 Ghz Intel Xeon® CPU with RAM 12.2Go and NVIDIA Tesla $c2050/c2070$ GPU computing processors, and running under Windows 7 64 bits. According to Figure 2.13, the computation time of our method is less time consuming than our early approach [28], because the description process relies on one measure

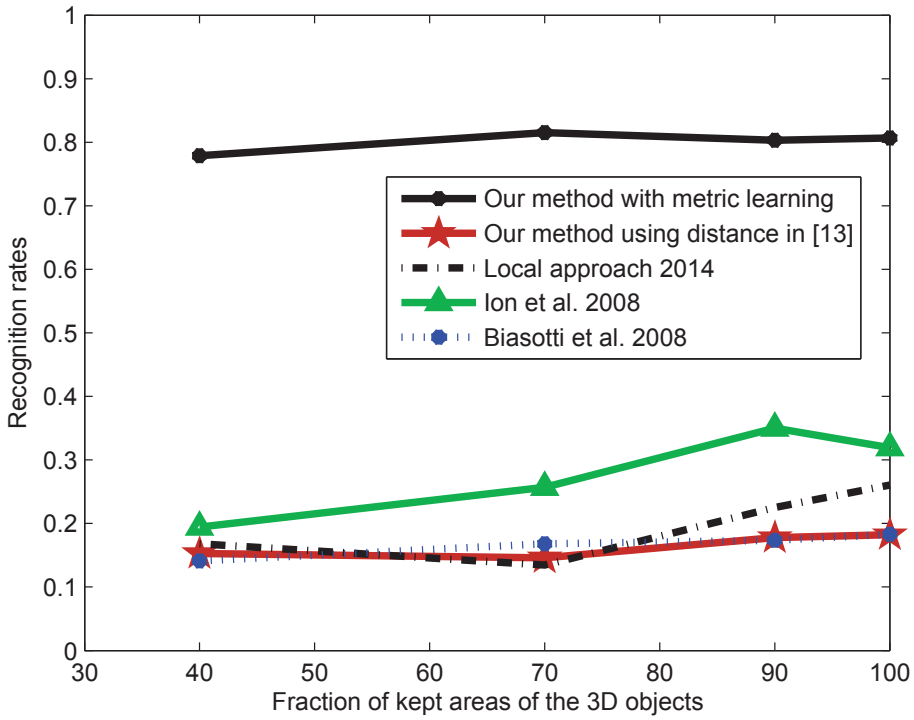

Figure 2.11 – The recognition rates for different methods according to the fraction of kept areas of $3D$ objects in the collection of Rodola et al. [54].

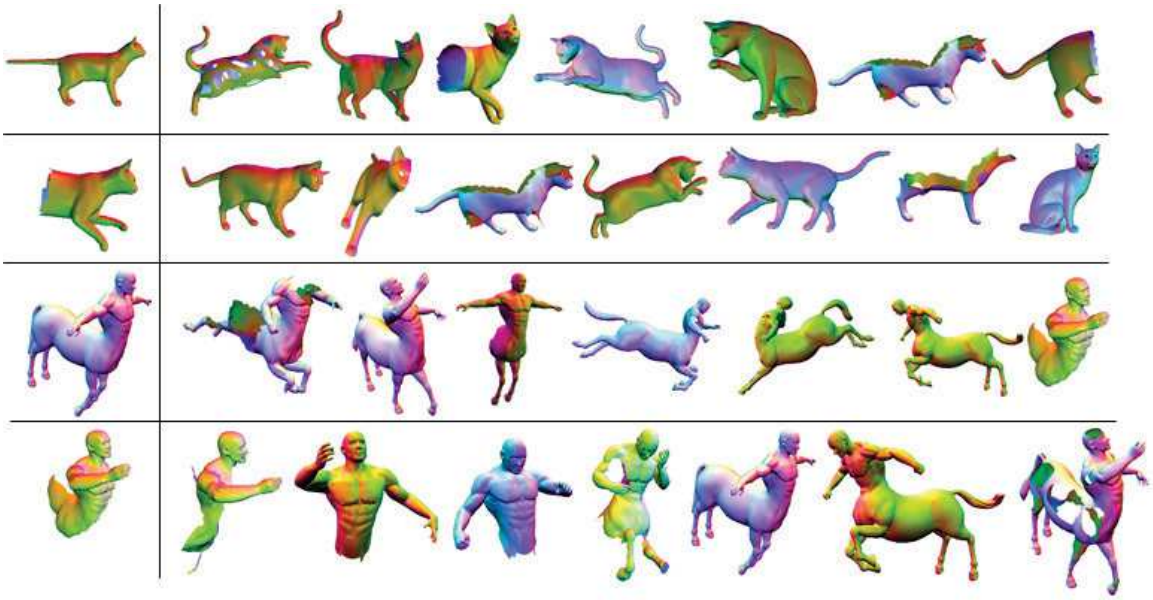

Figure 2.12 – Examples of queries $3D$ objects and retrieved $3D$ using cats and centaurs from the $3D$ objects collection proposed by Rodola et al. [54].

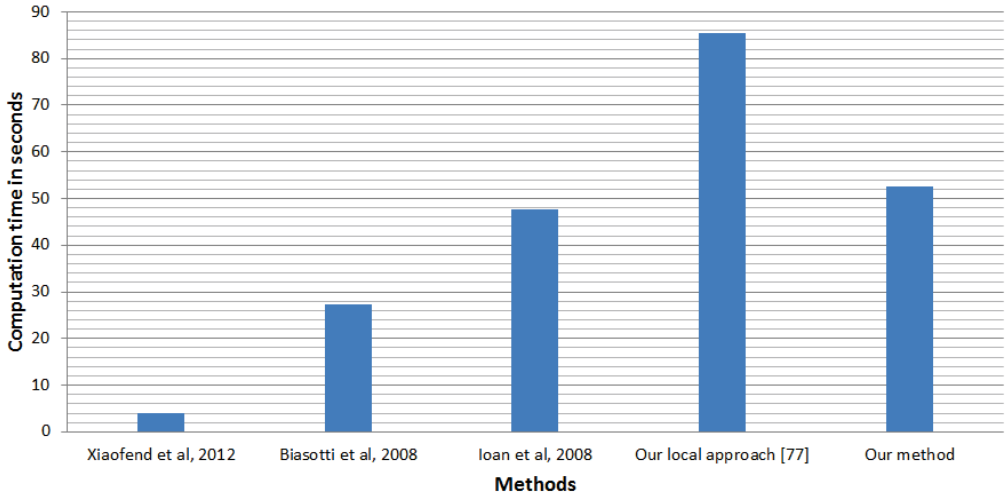

Figure 2.13 – Computational time of different methods.

Table 2.6 – The recognition rates for different methods according to the fraction of kept areas of 3D objects in the collection of Rodola et al. [54].

Fraction of kept areas	100%	90%	70%	40%
Our method with metric learning	**80.68%**	**80.31%**	**81.53%**	**77.88%**
Our method using distance in [13]	18.23%	17.77%	14.60%	15.28%
Local approach 2014	26.05%	22.50%	13.46%	16.81%
Ion et al. 2008	31.93%	35.01%	25.70%	19.42%
Biasotti et al. 2008	18.29%	17.35%	16.79%	14.11%

function. However, it is more time-consuming than the other approaches of the related work, because the computation of the measure function relies on the computation of the geodesic distances between vertices' pair of the 3D objects.

2.5 Conclusion and future works

In this paper, we developed a new method which describes the 3D object by a size function. The main idea of our method is the usage of the topological invariants provided by the critical points to describe the 3D object. For that purpose, the 3D object is represented by a size function resulting from a measure function that captures the surface displacement feature along the 3D object, and that encompasses invariance to affine transformations, articulations, and torsions. At first, the method was validated by using the 3D objects collection of McGill of 457 models. Afterwards, the method was validated by using an extended 3D objects collection of 2902 models. In order to tackle with the problem of the partial matching of 3D objects, our method was evaluated by using 275 3D objects from the 3D objects collection proposed by Rodola et al. [54]. Our method got the highest scores. However, our method is more computation time-consuming than the other approaches, because the computation of the measure function relies on the computation of the geodesic distances between vertices' pair of the 3D objects. To overcome this limitation, we are planning to parallelize the computation of the measure function and the size function on GPU in order to accelerate the computation time of our method.

Bibliography

[1] A. Ion, N.M. Artner, G. Peyré, S.B.L. Marmol, W.G. Kropatsch, L.D. Cohen. 3D Shape Matching by Geodesic Eccentricity. Proceeding of Computer Vision and Pattern Recognition Workshops, pp. 1-8, 2008.

[2] M. H. Mousa. Matching $3D$ objects using principle curvatures descriptors. Proceedings of 2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, pp. 453-457, 2011.

[3] J. Rabin, G. Peyré, L.D. Cohen. Geodesic Shape Retrieval via Optimal Mass Transport. Proceeding of the 11th European Conference on Computer Vision, Springer-Verlag, pp. 771-784, 2010.

[4] M. Ferri, S. Lombardini, C. Pallotti. Leukocyte classification by size functions. In Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, IEEE Computer Society Press, pp. 223-229, 1994.

[5] C. Xiaofeng, G. Chuanwei, Z. Xutang, J. Lijun. 3D Model Retrieval Based on Projected Area at Mesh Vertex. Third International Conference on Digital Manufacturing and Automation, IEEE International. pp. 1-4, 2012.

[6] L. De Floriani, U. Fugacci, F. Iuricich. Homological Shape Analysis Through Discrete Morse Theory. Perspectives in Shape Analysis, pp. 187-209, 2016.

[7] C. Heine, H. Leitte, M. Hlawitschka, F. Iuricich, L. De Floriani, G. Scheuermann, H. Hagen, C. Garth. A Survey of Topology-based Methods in Visualization. Computer Graphics Forum, 35 (3), pp. 643-667, 2016.

[8] P. Frosini, C. Landi. Size Theory as a Topological Tool for Computer Vision. Pattern Recognition and Image Analisys, 9(4), pp. 596-603, 1999.

[9] M. Handouyahia, D. Ziou, S. Wang. Sign language recognition using moment-based size functions. In: Proc. of Vision Interface 99, pp. 210-216, 1999.

[10] A. Verri, C. Uras. Metric-topological approach to shape representation and recognition. Image and Vision Computing. 14, pp.189–207, 1996.

[11] S. Baloch, H. Krim. Object recognition through topo-geometric shape models using error-tolerant subgraph isomorphisms. IEEE Transactions on Image Processing, 19 (5), pp. 1191-1200, 2010.

[12] S. Biasotti, A. Cerri, P. Frosini, D. Giorgi, C. Landi. Multidimensional Size Functions for Shape Comparison. Journal of Mathematical Imaging and Vision, Springer, 32 (2), pp. 161 - 179, 2008.

[13] M. D'Amico, P. Frosini, C. Landi, Natural pseudo-distance and optimal matching between reduced size functions, Acta Applicandae Mathematicae, 109 (2), pp. 527-554, 2010.

[14] J.B. Tenenbaum, V. de Silva, J. Langford. A global geometric framework for nonlinear dimensionality reduction. Science 290: 22, pp. 2319-2323, 2000.

[15] D.Y. Chen, M. Ouhyoung. A 3D model alignment and retrieval system. Proceeding of International Computer Symposium, 2002.

[16] S. Biasotti, L. De Floriani, B. Falcidieno, P. Frosini, D. Giorgi, C. Landi, L. Papaleo, M. Spagnuolo. Describing shapes by geometrical-topological properties of real functions. Computing Surveys, ACM, 40 (4), 2008.

[17] R.C. Veltkamp. Shape Matching: Similarity Measure and Algorithms. Proceedings Shape Modelling International, IEEE Press, pp. 188-197, 2001.

[18] B.J. Super. Learning Chance Probability Functions for Shape Retrieval or Classification. Proceedings of the IEEE Workshop on Learning in Computer Vision and Pattern Recognition, 2004.

[19] McGill 3D Shape Benchmark http://www.cim.mcgill.ca/ shape/benchMark/

[20] J.W. H. Tangelder, R.C. Veltkamp. A survey of content based 3D shape retrieval methods. Multimedia Tools and Applications, 39, pp. 441-471, 2008.

[21] B. Bustos, D. Keim, D. Saupe, T. Schreck. Content-based 3D object retrieval. IEEE Computer Graphics and Applications, IEEE Computer Society, 27 (4), pp. 22-27, 2007.

[22] M. Riedmiller, H. Braun, A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm, Proc. ICNN, 1993.

[23] M. D'Amico. A new optimal algorithm for computing size function of shapes. In CVPRIP Algorithms III: Proceedings of the International Conference on Computer Vision, Pattern Recognition and Image Processing, pp. 107-110, 2000.

[24] M. Minsky, S. Papert. Perceptrons: An Introduction to Computational Geometry. MIT Press. ISBN 0-262-63022-2, 1969.

[25] E.W. Dijkstra. A note on two problems in connexion with graphs. Numerische Mathematik, 1, pp. 269-271, 1959.

[26] D.L. Olson, D. Delen. Advanced Data Mining Techniques. Springer, 1st edition, 2008.

[27] P. Frosini, M. Pittore. New methods for reducing size graphs. Int. J. Comput. Math., 70 (3), pp. 505–517, 1999.

[28] M.A. Alaoui Mhamdi, D. Ziou. A local approach for 3D object recognition through a set of size functions. Image and Vision Computing, Volume 32 (12), pp. 1030-1044, 2014.

[29] J. Ponce, M. Hebert, C. Schmid, A. Zisserman (Eds.). Toward Category-Level Object Recognition. Lecture Notes in Computer Science 4170, ISBN 3-540-68794-7. Springer 2006.

[30] S. Dickinson, A. Leonardis, B. Schiele, M. Tarr, (Editors). Object Categorization: Computer and Human Vision Perspectives. Cambridge University Press, 2009.

[31] P. Chen, R. G. Steen, A. Yezzi, H. Krim. Joint brain parametric T_1-map segmentation and RF inhomogeneity calibration. Journal of Biomedical Imaging, 2009, pp. 1-14, 2009.

[32] U.Bagci, X. Chen, J.K.Udupa. Hierarchical Scale-Based Multi-Object Recognition of 3D Anatomical Structures. IEEE Transactions on Medical Imaging, 31 (3), 2012.

[33] Y. Takahashi, Y. Muramatu, K. Kato. Object Recognition using Simultaneous Pattern-Matching in Virtual Environment. SICE-ICASE International Joint Conference, pp. 5579-5583, 2006.

[34] Y. Lei, Y. Guo, M. Hayat, M. Bennamoun, X. Zhou. A Two-Phase Weighted Collaborative Representation for 3D partial face recognition with single sample. Pattern Recognition, 52, pp. 218-237, 2016.

[35] D. Lee, H. Krim. $3D$ face recognition in the Fourier domain using deformed circular curves. Multidimensional systems and signal processing, 28 (1), pp. 105-127, 2017.

[36] F. Stanco, S. Battiato. Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks. CRC Press, 2011.

[37] F. Bosché. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. Advanced Engineering Informatics, 24 (1), 107-118, 2010.

[38] M. Wang, Y. Gao, K. Lu, and Y. Rui. View-Based Discriminative Probabilistic modeling for 3D Object Retrieval and Recognition. IEEE Transactions on Image Processing, 22 (4), pp. 1395-1407, 2013.

[39] A. Verri, C. Uras, P. Frosini, M. Ferri. On the use of size functions for shape analysis. Biolog. Cybernet. 70, pp. 99-107, 1993.

[40] H. Tabia, H. Laga. Covariance-based Descriptors for efficient 3D shape matching, retrieval and classification. IEEE Transactions on Multimedia, 17 (9), pp. 1591-1603, 2015.

[41] A. Agathos, I. Pratikakis, P. Papadakis, S. Perantonis, P. Azariadis, N. Sapidis. Retrieval of 3D articulated objects using a graph-based representation. ACM 3DOR, pp. 29-36, 2009.

[42] H. Tabia, D. Picard, H. Laga, P.-H. Gosselin. Compact vectors of locally aggregated tensors for 3D shape retrieval. Proc. ACM 3DOR, pp. 17-24, 2013.

[43] G. Lavoué. Combination of bag-of-words descriptors for robust partial shape retrieval. Vis. Comput., 28 (9), pp. 931-942, 2012.

[44] Robert Sumner's shapes repository http://people.csail.mit.edu/ sumner/research/deftransfer/

[45] M. D'Amico, M. Ferri, I. Stanganelli. Qualitative Asymmetry Measure for Melanoma Detection. IEEE International Symposium on Biomedical Imaging , pp. 1155-1158, 2004.

[46] M. Dutton, R. Doran. Autodesk 3ds Max 8 Revealed. Course Technology Inc, 2005.

[47] Z. Lian, A. Godil, P. Rosin, X. Sun. A new convexity measurement for 3d meshes. IEEE CVPR, pp. 119-126, 2012.

[48] R. Baeza-Yates, G. Valiente. An Image Similarity Measure Based on Graph Matching. Proceedings of the Seventh International Symposium on String Processing Information Retrieval, pp. 28-38, 2000.

[49] S. Suresh, N. Sundararajan, P. Saratchandran. Risk-sensitive loss functions for sparse multi-category classification problems. Information Sciences, 178 (12), pp. 2621-2638, 2008.

[50] P. Zezula, G. Amato, V. Dohnal, M. Batko. Similarity Search: The Metric Space Approach. The Kluwer International Series on Advances in Database Systems, Springer, Volume 32, 2006.

[51] J. S. Beis and D. G. Lowe. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1000-1006, 1997.

[52] J.J. Freeman. Experiments in discrimination and classification. Pattern Recognition, 1(3), pp. 207-218, 1969.

[53] P. Papadakis, I. Pratikakis, T. Theoharis, G. Passalis, and S. Perantonis. 3D object retrieval using an efficient and compact hybrid shape descriptor. Proc. ACM 3DOR, pp. 1-7, 2008.

[54] E. Rodola, L. Cosmo, M. M. Bronstein, A. Torsello, D. Cremers. Partial Functional Correspondence. Computer Graphics Forum, 36 (1), pp. 222-236, 2017.

[55] D. V. Vranic, D. Saupe. $3D$ Shape Descriptor Based on 3D Fourier Transform. EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS 01), pp. 271-274, 2001.

[56] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller. Multi-view Convolutional Neural Networks for 3D Shape Recognition. Proceedings of ICCV, 2015.

[57] H. Zimmer, M. Campen, L. Kobbelt. Efficient Computation of Shortest Path-Concavity for 3D Meshes. IEEE CVPR, pp. 2155-2162, 2013.

[58] P. Frosini. Measuring shapes by size functions. In Intelligent Robots and Computer Vision X: Algorithms and Techniques, 1607, pp. 122-133, 1991.

[59] H. Laga, H. Takahashi, M. Nakajima. Spherical Wavelet Descriptors for Content-based 3D Model Retrieval. IEEE International Conference on Shape Modeling and Applications (SMI2006), pp. 75-85, 2006.

[60] A. Arampatzis, J. Kamps, S. Robertson. Where to stop reading a ranked list?: threshold optimization using truncated score distributions. Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 524-531, 2009.

[61] C. Villani. Topics in Optimal Transportation. American Mathematical Society, 2003.

[62] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser. The Princeton Shape Benchmark. In the proceedings of Shape Modeling International, pp. 167-178, 2004.

[63] M. Attene. A lightweight approach to repairing digitized polygon meshes. International Journal of Computer Graphics, 26 (11), pp. 1393-1406, 2010.

[64] R. Wessel, I. Blümel, R. Klein. A 3D Shape Benchmark for Retrieval and Automatic Classification of Architectural Data. In proceedings of Eurographics 2009, Workshop on 3D Object Retrieval, pp. 53-56, 2009.

Conclusion

Dans cette thèse, nous nous sommes intéressés au problème de la reconnaissance d'objets $3D$. Pour cela, nous avons choisi de concevoir avec des méthodes de description d'objets $3D$ basées sur les informations fournies par les points critiques à travers la fonction de taille. Afin de justifier le choix de la méthode de reconnaissance d'objets $3D$ à concevoir, dans chacune des deux contributions réalisées, nous avons fait une analyse des approches de reconnaissance d'objets $3D$ dans la littérature. De ce fait, nos contributions peuvent être récapitulées comme suit.

Dans le premier chapitre, nous avons proposé une approche locale de la reconnaissance d'objets $3D$ à travers un ensemble de 18 fonctions de taille. D'abord, afin de doter la méthode de description de l'invariance aux articulations, torsions et transformation affines, un processus de normalisation est effectué en deux étapes. La première étape consiste à utiliser la projection ISOMAP pour que les objets $3D$ de la même classe aient les même torsions et articulations. Dans la deuxième étape, l'ACP est utilisé pour assurer l'invariance aux translations, échelles et rotations. Ensuite, chaque objet $3D$ est découpé en 18 portions afin de représenter chacune par une fonction de taille résultante d'une fonction de mesure, qui représente les caractéristiques de concavité et convexité de chaque vertex de l'objet $3D$. Enfin, pour comparer les objets $3D$, une distance qui exprime la moyenne ramifiés de minimums entre les fonctions de taille associés aux objets $3D$ est utilisée pour mesurer la similarité entre les objets $3D$. Les scores obtenus expriment que notre approche locale peut être utilisée pour la reconnaissance d'objets $3D$ ayant subi différentes transformations affines, et qui ont des articulations et des torsions différentes. Ce travail été publié dans un articule intitulé **A local approach for** $3D$ **object recognition through a set of size functions**, dans la revue **Image and Vision Computing, Elsevier** en 2014.

Dans le deuxième chapitre, nous avons proposé une approche globale de la reconnaissance d'objets $3D$ à travers une seule fonction de taille. Le but est de pallier aux limitations de la première contribution et d'entreprendre le problème de la reconnaissance partielle d'objets $3D$. Pour cela, l'objet $3D$ est représenté par une fonction de taille résultante d'une fonction de mesure qui exprime la caractéristique du déplacement surfacique que nous avons conçue. En effet, nous avons pu incorporer l'invariance aux transformations affines, articulations et torsions dans la fonction de mesure pour éviter le processus de la normalisation d'objet $3D$. Afin d'adapter notre méthode pour qu'elle soit utilisée dans le cadre de la reconnaissance partielle d'objets $3D$, nous avons utilisé un réseau de neurones de quatre couches pour pondérer les appariements entre les points ou lignes coins des fonctions de taille associées aux objets $3D$ selon leurs pertinences. Les résultats obtenus expriment que notre méthode peut être utilisée pour la reconnaissance d'objets $3D$ complets et objets $3D$ avec une ou plusieurs portions manquantes. Cette contribution est en cours de révision dans un article intitulé $3D$ **object recognition through a size function resulting from an invariant topological feature** dans la revue **IEEE Transactions on Pattern Analysis and Machine Intelligence**.

Dans cette thèse, la description d'objets $3D$ est basée sur les informations fournies par les points critiques à travers la fonction de taille. Comme déjà décrit dans les deux articles, c'est que ces informations sont les composantes connexes entre les points critiques de l'objets $3D$ et les relations entre eux. Le choix des fonctions de mesure effectué nous a permis d'aboutir à des scores qui dépassent ceux de l'état de l'art. Afin d'améliorer davantage les résultats obtenus, nous souhaitons ajouter les informations sur les cavités de l'objet $3D$. Nous planifions l'apprentissage de la fonction de mesure et une reformulation probabiliste des distances de correspondance.

Au-delà de l'avancement de la connaissance, les contributions réalisées pourraient avoir d'importantes retombées socio-économiques et socioculturelles. En effet, elles peuvent être utilisée dans le cadre de la sécurité comme la reconnaissance faciale en $3D$ et la détection d'objets suspicieux. Nous planifions l'orientation les contributions dans le domaine de la robotique pour la localisation et la cartographie simultanées d'objets $3D$ dans les grandes scènes, et la conduite autonome de véhicules.

Bibliographie

[1] R. C. Gonzalez, R. E. Woods. Digital Image Processing, Fourth Edition. Pearson, 2017.

[2] E. R. Davies. Computer Vision : Principles, Algorithms, Applications, Learning, Fifth Edition. Academic Press, 2017.

[3] J. Liu, B.Ye, J. Peng, B. Wu. Fundamentals of Computer Graphics, Fourth Edition. A K Peters/CRC Press, 2015.

[4] N. Magnenat-Thalmann, D. Thalmann . Image Synthesis : Theory and Practice. Springer, 2011.

[5] G. Ferrand, J. English, P. Irani. $3D$ Visualization of Astronomy Data Cubes using Immersive Displays. Annual General Meeting of the Canadian Astronomical Society, 2016.

[6] P. Stelldinger, L.J. Latecki. $3D$ Object Digitization : Topology Preserving Reconstruction. Proceedings of the IEEE International Conference on Image Analysis, 2006.

[7] P. R. Cromwell. Polyhedra. Cambridge University Press, New York, 1997.

[8] L. Wei L, A. Sourin, O. Sourina. Function-based Visualization and Haptic Rendering in Shared Virtual Spaces, The Visual Computer, 24(10), pp. 871-880, 2008.

[9] J.W.H. Tangelder, R.C. Veltkamp. A survey of content based $3D$ shape retrieval methods. Multimedia Tools and Applications, 39, pp. 441-471, 2008.

[10] P. Alliez, C. Gotsman. Recent advances in compression of $3D$ meshes. In Advances in Multiresolution for Geometric Modelling (ed. Dodgson N, Floater M and Sabin M), Springer-Verlag, pp. 3-26, 2005.

[11] P. Alliez, M. Desbrun. Progressive encoding for lossless transmission of $3D$ meshes. ACM SIGGRAPH, pp. 198–205, 2001.

[12] A. Bors. Blind watermarking of $3D$ shapes using localized constraints. IEEE 2nd International Symposium on $3D$ Data Processing, Visualization and Transmission, pp. 242–249, 2004.

[13] D. Koller, M. Levoy. Computer-aided Reconstruction and New Matches in the Forma Urbis Romae. Proc. Formae Urbis Romae - Nuove Scoperte, Bullettino Della Commissione Archeologica Comunale di Roma, 2006.

[14] O. Villar. Learning Blender : A Hands-On Guide to Creating $3D$ Animated Characters, first edition. Addison-Wesley Professional, 2014.

[15] D. Eberly. $3D$ Game Engine Design : A Practical Approach to Real-Time Computer Graphics (Morgan Kaufmann Series in Interactive $3D$ Technology), Second Edition. CRC Press. 2006.

[16] G. Sellers, R.S. Wright Jr., N. Haemel. OpenGL Superbible : Comprehensive Tutorial and Reference (7th Edition). Addison-Wesley Professional, 2015.

[17] H. Zhang, D. Liang. Computer Graphics Using Java $2D$ and $3D$. Prentice Hall, first edition, 2006.

[18] S. Biasotti , L. De Floriani, B. Falcidieno, P. Frosini, D. Giorgi, C. Landi, L. Papaleo, M. Spagnuolo. Describing shapes by geometrical-topological properties of real functions. ACM Computing Surveys, 40(4), 2008.

[19] E. Rodola, L. Cosmo, M. M. Bronstein, A. Torsello, D. Cremers. Partial Functional Correspondence. Computer Graphics Forum, 36 (1), pp. 222-236, 2017.

[20] Marc Levoy, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, Sean Anderson, James Davis, Jeremy Ginsberg, Jonathan Shade, and Duane Fulk. The Digital Michelangelo Project : $3D$ Scanning of Large Statues. Proceedings of ACM SIGGRAPH, pp. 131–144, 2000.

[21] D. Gallup, J.-M. Frahm, M. Pollefeys. Piecewise Planar and Non-Planar Stereo for Urban Scene Reconstruction. IEEE Conference Computer Vision and Pattern Recognition, pp. 1418-1425, 2010.

[22] Szymon Rusinkiewicz. Estimating Curvatures and Their Derivatives on Triangle Meshes. Symposium on $3D$ Data Processing, Visualization, and Transmission, 2004.

[23] J.A. De Loera, J. Rambau, F. Santos. Triangulations : Structures for Algorithms and Applications. Springer-Verlag, 2010.

[24] G. T. Herman. Fundamentals of computerized tomography : Image reconstruction from projection, 2nd edition. Springer, 2009.

[25] F. Lecron, M. Benjelloun, S. Mahmoudi. Descriptive Image Feature for Object Detection in Medical Images. International Conference Image Analysis and Recognition, pp. 331-338, 2012.

[26] L. Carrer, A. G. Yarovoy. Concealed weapon detection using UWB 3-D radar imaging and automatic target recognition. The 8th European Conference on Antennas and Propagation, 2014.

[27] M. Wertheimer. Untersuchungen zur Lehre von der Gestalt, II [Laws of organization in perceptual forms]. Psychologische Forschung, 4, pp. 301-350, 1923.

[28] W. D. Ellis. A source book of Gestalt psychology. London : K. Paul, Trench, Trubner & Co, 1938.

[29] J. Ponce, M. Hebert, C. Schmid, A. Zisserman (Eds.). Toward Category-Level Object Recognition. Lecture Notes in Computer Science 4170, ISBN 3-540-68794-7, 2006.

[30] S. Dickinson, A. Leonardis, B. Schiele, and M. Tarr, (Eds). Object Categorization : Computer and Human Vision Perspectives. Cambridge University Press, 2009.

[31] C.B. Mervis, E.H. Rosch. Categorization of Natural Objects. Annual Review of Psychology, 32, pp. 89-113, 1981.

[32] A. Oliva, A. Torralba. The role of context in object recognition. Trends in Cognitive Sciences, vol. 11(12), pp. 520-527, 2007.

[33] M. J. Choi, A. Torralba, A. S. Willsky. Context Models and Out-of-context Objects. Pattern Recognition Letters, 33(7), pp. 853-862, 2012.

[34] I. Biederman, R.J. Mezzanotte, J. C. Rabinowitz. Scene perception : Detecting and judging objects undergoing relational violations. Cognitive Psychology, 14(2), pp. 143-177, 1982.

[35] A. Oliva, A. Torralba. The role of context in object recognition. Trends in Cognitive Sciences, 11(12), pp. 520-527, 2007.

[36] D. Lee, H. Krim. $3D$ surface reconstruction using structured circular light patterns.International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 279-289, 2010.

[37] D. Lee, H. Krim. $3D$ face recognition in the Fourier domain using deformed circular curves. Multidimensional systems and signal processing, 28(1), pp. 105-127, 2017.

[38] S. P. Johnson. How Infants Learn About the Visual World. Cognitive Science, 34, pp. 1158-1184, 2010.

[39] D. G. Pelli, N. J. Majaj, N. Raizman, C. J. Christian, E. Kim, M. C. Palomares. Grouping in object recognition : the role of a Gestalt law in letter identification. Cognitive Neuropsychology, 26(1), pp. 36-49, 2009.

[40] E. Michaelsen, D. Muench, M. Arens. Recognition of Symmetry Structure by Use of Gestalt Algebra. IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 206-210, 2013.

[41] A. F. Pereira, L. B. Smith. Developmental changes in visual object recognition between 18 and 24 months of age. Developmental Science, 12(4), 67-80, 2009.

[42] L. Zhu. Computer Model of Geological Faults in $3D$ and the Application in Beijing Olympic Green District. 16th International Conference on Artificial Reality and Telexistence–Workshops, 2006.

[43] P.B. Xu, H. Xu, H.M. Wen. $3D$ meso-mechanical modeling of concrete spall tests. International Journal of Impact Engineering, 97, pp. 46-56, 2016.

[44] J. Liu, B.Ye, J. Peng, B. Wu. New Method of Feature Reconstruction from Engineering Drawings Based on Similarity Measure of Primitive Views. Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009.

[45] A.F. Ayoub, Y. Xiao, B. Khambay, J.P. Siebert, D. Hadley. Towards building a photo-realistic virtual human face for craniomaxillofacial diagnosis and treatment planning. International Journal of Oral and Maxillofacial Surgery, 5, pp. 423–428. 2007.

[46] L. De Floriani, U. Fugacci, F. Iuricich. Homological Shape Analysis Through Discrete Morse Theory. Perspectives in Shape Analysis, pp. 187-209, 2016.

[47] C. Heine, H. Leitte, M. Hlawitschka, F. Iuricich, L. De Floriani, G. Scheuermann, H. Hagen, C. Garth. A Survey of Topology-based Methods in Visualization. Computer Graphics Forum, 35 (3), pp. 643-667, 2016.